AFHRL-TR-89-69

AIR FORCE

HUMAN RESOURCES

AD-A221 744

AUTOMATIC INFORMATION PROCESSING AND HIGH PERFORMANCE SKILLS: ACQUISITION, TRANSFER, AND RETENTION

Arthur D. Fisk
Kevin A. Hodge
Mark D. Lee
Wendy A. Rogers

Georgia Institute of Technology
School of Psychology
Atlanta, Georgia 30332

LOGISTICS AND HUMAN FACTORS DIVISION
Wright-Patterson Air Force Base, Ohio 45433-6503

April 1990
Interim Technical Report for Period August 1988 – October 1989

LABORATORY

AIR FORCE SYSTEMS COMMAND
BROOKS AIR FORCE BASE, TEXAS 78235-5601

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Public Affairs Office has reviewed this report, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This report has been reviewed and is approved for publication.

BERTRAM W. CREAM, Technical Director
Logistics and Human Factors Division

HAROLD G. JENSEN, Colonel, USAF
Commander

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (*Leave blank*)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	April 1990	Interim - August 1988 to October 1989

4. TITLE AND SUBTITLE
Automatic Information Processing and High Performance Skills: Acquisition, Transfer, and Retention

5. FUNDING NUMBERS
C - F33615-88-C-0015
PE - 62205F
PR - ILIR
TA - 40
WU - 01

6. AUTHOR(S)
Arthur D. Fisk, Mark D. Lee
Kevin A. Hodge, Wendy A. Rogers

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Georgia Institute of Technology
School of Psychology
Atlanta, Georgia 30332

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Logistics and Human Factors Division
Air Force Human Resources Laboratory
Wright-Patterson Air Force Base, Ohio 45433-6503

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
AFHRL-TR-89-69

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (*Maximum 200 words*)

Ten experiments involving basic laboratory research on automatic processing theory and skill acquisition are reviewed. The experiments were conducted to investigate the following issues: effects of modified practice, transfer of training, skill decay, and retention. The results of this work provide an understanding of skill acquisition, retention, and transfer with respect to high performance skills training.

14. SUBJECT TERMS
automaticity, skill retention, part-task training, skill transfer, skill acquisition, training

15. NUMBER OF PAGES
144

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

SUMMARY

This document summarizes a basic research effort investigating automatic processing theory and high performance skills training. Research issues such as skill acquisition, skill retention and transfer of training are explored. The results of this work suggest that the application of automatic processing theory to training complex skills can have an impact on skill acquisition.

PREFACE

The work documented in this report was conducted under Air Force Human Resources Laboratory (AFHRL) Contract No. F33615-88-C-0015 with the University of Dayton Research Institute and was performed by the subcontractor Georgia Institute of Technology Research Institute. This work supports an integrated research program which is developing advanced part-task training techniques based on information processing theory. Captain Michael T. Lawless served as the AFHRL/LRG, Wright-Patterson AFB, contract monitor.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

AUTOMATIC INFORMATION PROCESSING AND HIGH PERFORMANCE SKILLS: ACQUISITION, TRANSFER, AND RETENTION

I. OVERVIEW OF THE EXPERIMENTAL INVESTIGATION

This document details four series of experiments (a total of 10 individual experiments) that were conducted to further extend automatic/controlled processing research to command and control mission training. The purpose of the present program of research was to investigate training-program-relevant issues that had not previously been addressed in the literature. These issues can be categorized as (a) acquisition, (b) transfer, and (c) retention of high performance skilled behavior. Transfer and retention in the realm of automatic processing had not been extensively investigated prior to the present research effort but are issues clearly important for understanding how to maximize operator training and how to maintain mission readiness over extended time periods.

This document describes experiments that examine multiple facets of component training: (a) what transfers, (b) the limits of transfer, (c) retention of task component skills, (d) the important issue of component recombination and how to combine component training for maximum training benefit (in terms of minimum training time and costs). Because of the breadth of the issues examined, each of four primary, independent sections exhaustively investigates one aspect of the above research domains.

The first series of experiments investigated the effects of type and amount of consistent mapping practice on automatic process development. These experiments begin the investigation of the effects of differential amounts of practice on the "strength" (degree of automatic process development) of consistently mapped stimulus items. These experiments help to assess when it is possible to reduce the amount of practice needed for a given level of skill development. To briefly summarize the findings from this series, the data confirm that, in general, the more consistent mapping practice a person receives, the better his/her performance will be at the end of the training. More important, the data suggest that it may be possible to specify how to combine training such that some training elements will benefit from the training of other elements; hence, training time can be reduced. If a "superset" can be formed during training (and that set can be formed quickly), then detection of one stimulus item seems to strengthen the entire to-be-trained set.

The second series of experiments examined transfer of training as a function of semantic relatedness in a task assessing performance at a subject's perceptual processing limits. Three experiments were conducted. Subjects participated in a semantic category search task and we used an adaptive-training, multiple-frame paradigm. This was a newly developed paradigm that provided progressively less processing time as a given subject increased his or her detection accuracy. The paradigm allowed us to assess performance at a subject's individual perceptual processing limit. After practice on various semantic categories, subjects attempted to detect words which they had not been explicitly trained on but were words from the trained category, from a highly related category (related relative to the trained category), from a moderately related category or from an unrelated category. Semantic transfer was not perfect but it was impressive given that the dominant task components were visual in nature. We found positive transfer for the untrained exemplars from the trained category and the pattern of performance of the other conditions was lawfully related to the degree of semantic relatedness. We also found that, for this class of tasks at least, the more training a subject received, the greater the semantic transfer.

The third experimental series was conducted to investigate the transfer of automatic component processes to compatible, incompatible, and conflict situations. The issue investigated related mostly to retraining. This investigation provided a systematic examination of transfer of automatic processes to situations where the fully trained automatic components were used in the same way, in an opposite manner, or in conflict with other automatic processes. Subjects received extensive semantic category search practice and then were transferred to the above-mentioned situations. At the first session of transfer, we found positive transfer when automatic components were used in the same manner, and negative transfer when automatic components were used in an opposite manner. The most striking results are from the conflict situations. There was substantial disruption when target stimuli were used in conflict with other previously trained target stimuli. However, there was much less disruption, and it diminished quickly, for situations where distractor stimuli were used in conflict with other previously trained distractor stimuli. These data demonstrate the need to incorporate both transfer and disruption functions when designing training programs. The study emphasizes the need to understand consistent components of tasks for specifying potential conflict situations both within the same task and for related training tasks.

The final series of experiments examined long-term retention relative to automatic component processes. Clearly this issue is important because situations exist where personnel are trained and then use the skill only when an emergency arises. Given this kind of scenario, we need to be able to predict the mission readiness of trainees. We also need information to predict the timeframe and the potential need for refresher training. This series of experiments gives us this information, at least for the class of tasks used herein. In this series, we investigated very long-term retention, retention up 180 days after training. For that long-term retention experiment, we utilized a hybrid memory/visual search task. We conducted two other experiments in an attempt to more fully isolate the locus of performance decay by examining separate retention functions for pure memory search and for pure visual search. In the first experiment, we found very little decay (13 percent maximum) over the 180 days. We found that the maximum decay occurred within 30 days subsequent to training. The component approach to understanding the decay of skilled performance demonstrated very little automatic, direct-memory-access decay. We found no decay in semantic access and some decay (but minimal) in visual search processes. We concluded that the decay occurs in the control structure such as strategic processing or interaction among component processes.

In the following detailed account of the experimental investigation, each section is self-contained so that the reader interested in only some of the issues can turn immediately to the relevant section(s).

II. EXPERIMENTAL SERIES 1: EFFECTS OF TYPE AND AMOUNT OF CONSISTENT MAPPING PRACTICE

Introduction

Optimization of final-level performance is the goal of most training programs designed to aid the acquisition of skilled behavior. There are many variables, however, which affect the ultimate success of training programs. For example, what type of practice should be provided? How much practice? Is more practice always better? Is there a point at which more practice will not yield substantial performance improvements? Is it better to practice on groups of similar tasks or subtasks, or is it better to distribute practice (to some degree) across similar components of tasks? The goal of the present series of experiments was to begin to answer some of these questions.

Using a visual search paradigm, we investigated the differences between consistent and inconsistent practice in terms of the amount of overall training both between and within subjects, as well as between and within blocks of trials. We were interested in the effects of these variables on the development of automatic processing, an integral component of most skilled behavior (Logan, 1985). Automatic processing has been investigated extensively in the memory search domain, and to a lesser degree in the visual search domain (e.g., Schneider & Shiffrin, 1977). We first review automatic and controlled processing theory and the differential performance characteristics that result from consistent versus inconsistent, or variable practice. Following this review is a brief description of a strength-based theory of the development of automatic processing. The present series of experiments made use of these fundamental theoretical and empirical foundations in an effort to examine the various factors which contribute to the success of training programs designed to improve skill development.

Automatic and Controlled Processes

A well-documented finding in the realm of attention research is that there are two qualitatively different types of information processing which interact in the performance of most complex tasks (LaBerge & Samuels, 1974; Logan, 1978, 1979, 1985, 1988a, 1988b; Posner & Snyder, 1975; Schneider, Dumais, & Shiffrin, 1984; Schneider & Shiffrin, 1977; Shiffrin, 1988; Shiffrin & Dumais, 1981; Shiffrin & Schneider, 1977). Following the lead of Schneider and Shiffrin (1977), these two processes will be referred to herein as "automatic" and "controlled" processes.

Automatic processes are characterized as fast, parallel, fairly effortless, and not limited by short-term memory capacity; these processes are difficult to acquire and, once well learned, difficult to modify. Furthermore, automatic prccesses are not sensitive to vigilance decrements (Fisk & Schneider, 1981), alcohol intoxication (Fisk & Schneider, 1982), fatigue (Hancock, 1984), or heat stress (Hancock & Pierce, 1984).

Controlled processes, on the other hand, are generally slow, serial, attention-demanding, and limited by short-term memory capacity. (For a more detailed analysis of the characteristics of automatic and controlled processing see Fisk, Ackerman, &

Schneider, 1987; Logan, 1985; Posner & Snyder, 1975; Schneider, Dumais, & Shiffrin, 1984; Shiffrin, 1988; Shiffrin & Dumais, 1981.)

Controlled processing components usually dominate in the performance of novel tasks. However, if major components of the task are consistent, performance can become automatized after substantial practice. A central goal of training research is to understand how, and under what conditions, performance improves. Generally speaking, an important component of many training programs involves training the consistent elements of a task (Schneider, 1985a).

In their series of experiments investigating controlled search and automatic detection, Schneider and Shiffrin (1977; Shiffrin & Schneider, 1977) demonstrated differences in performance as a function of whether training was consistent or varied. The degree of consistency in the relationship between the stimulus (or classes of stimuli) and the response requirements has been referred to as consistent or varied "mapping." More precisely, in a consistent mapping (CM) situation the individual always deals with (i.e., attends to or responds to or utilizes information from) a stimulus, or class of stimuli, in the same manner. CM training conditions result in dramatic performance improvements (see Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977 for details), modifications in the characteristics of event-related brain potentials (Kramer, Schneider, Fisk, & Donchin, 1986), and the eventual development of performance characteristics indicative of automatic processing. Varied mapping (VM) training situations are those in which the practice is inconsistent; that is, the response or degree of attention devoted to the stimulus changes from one stimulus exposure to another. VM training conditions result in relatively little performance improvement.

Automatic Process Development

Many theories of automatic process development are based on the modal view of a strength representation of knowledge (e.g., Anderson, 1982, 1983; Dumais, 1979; LaBerge & Samuels, 1974; MacKay, 1982; Schneider, 1985b; Schneider & Detweiler, 1987, 1988; Shiffrin & Czerwinski, 1988; but see Logan, 1988a, 1988b, for a non-strength theory). All of these theories propose that some increase and/or decrease in "strength" is responsible for the development of automaticity.

The concept of strength varies among the models, but is generally related to the role or significance of a stimulus or set of stimuli, a rule, or a connection (e.g., between nodes). For example, MacKay's (1982) strength theory is based on repeated activation, priming, reinforcement, and the resultant changes in strength between nodes. Production system models incorporate a conceptualization of strength associated with production rules. Strength is increased when a rule is invoked and weakened when application of the rule leads to error. According to Neches, Langley, and Klahr (1987), "the strength (or weight) of a production is a parameter that is adjusted to indicate the system's current confidence in the correctness and/or usefulness of that rule" (p. 39). Finally, connection system models are all strength-based, in that they assume that knowledge is the strength of connections among units of information (for a review, see Rumelhart & McClelland, 1987).

Recently, Schneider (Schneider, 1985b; Schneider & Detweiler, 1987, 1988) has proposed an eclectic strength model which is a hybrid of production system and connectionist models. According to Schneider's connectionist/control model, the development of automaticity is a function of two types of learning mechanisms: associative and priority learning, both of which are strength-based.

The associative learning mechanism alters the connection weights between input and output information so that, after sufficient training, a given input comes to evoke the associated output. Furthermore, associative learning results in the strengthening of connections between stimuli (e.g., members of a category) so that activation of one stimulus results in the activation of others.

The priority learning mechanism modifies how strongly a given message (i.e., stimulus information) is transmitted. This strength of transmission is defined as the "priority tag" of that message. A key element of priority learning is that the increment or decrement of a priority tag is based on whether or not a message is important; that is, whether or not prior presentation of that message produced a substantial amount of subsequent processing. Important messages have high priority tags and unimportant messages have low priority tags.

It is assumed that consistent practice leads to continual incrementing of the priority tag for target stimuli (when detected) and decrementing of the priority for distractor stimuli. Thus, CM practice leads to a segregation of stimuli so that stimuli with high priority tags (consistent targets) become "foreground" and stimuli with very low priority tags (consistent distractors) become "background." Within Schneider's hybrid connectionist model, pure automatic processing (processing without control process assistance) is not possible without sufficient priority learning. A combination of both associative and priority learning allows stimuli to be filtered and messages transmitted without control processing assistance; hence, stimuli can automatically attract attention. A common example of the presence of some stimulus or configuration of stimuli resulting in the automatic attraction of attention is the cocktail party phenomenon. This phenomenon is exemplified by the situation in which a person is listening to one conversation amid a din of background conversation yet attention is immediately drawn to another conversation when the person hears his or her own name.

Support for Strength Theory

Many experiments have provided evidence in support of the assumption that search performance is determined by the strength of the target relative to the strength of the distractor (e.g., Dumais, 1979; Prinz, 1979). On the first trial of training, it is assumed that all stimuli have an equivalent, intermediate strength (Dumais, 1979; Shiffrin & Czerwinski, 1988; Shiffrin & Dumais, 1981). The strength of the stimuli is intermediate and not zero because the stimuli are not completely novel but are simply untrained. For example, if words or letters are used as stimuli, they are familiar but have not been previously trained to have a high strength level, at least within the experimental context (Schneider & Fisk, 1984).

By definition, each time a CM target appears in the display it is always attended and/or responded to (except, of course, in the case of a "miss"). In this manner, the importance of a CM stimulus is increased and thus the CM stimulus becomes associated with a high priority tag. After many trials of CM training, the high priority associated with CM targets will result in these items being transmitted without the need for serial search. Consistent distractors, on the other hand, will have a decreased strength level after practice because their appearance results in either a negative response (e.g., correct rejection) or no response at all. Therefore, CM distractors will have a very low priority. Finally, VM stimuli maintain an intermediate strength because on some trials they are targets and are attended to, while on other trials they serve as distractors and must be ignored. Conceptually, the priority tag of the VM stimuli increases on some trials and decreases on other trials; therefore, even after many trials of training, these stimuli will still have an intermediate strength level.

Transfer and/or reversal of CM-trained targets and distractors yields a pattern of results which supports strength-based theories of perceptual learning. For example, Rabbitt, Cumming, and Vyas (1979) found that positive transfer (i.e., no disruption in performance) occurs when previously trained CM targets are paired with new distractor stimuli. According to a strength model, this is to be expected because targets which have been previously trained as CM targets have a higher strength relative to the novel stimuli used as distractors in the transfer condition. (As mentioned previously, novel stimuli have an intermediate strength level prior to training.)

Kristofferson (1977) demonstrated that positive transfer is also found when new targets are paired with previously trained CM distractors. In this case the CM distractors have a low strength level relative to the novel stimuli being used as targets. Although a strength theory is not explicitly formulated by Rabbitt or Kristofferson, their data provide evidence for both target learning and distractor learning in search tasks.

Dumais (1979) conducted a series of experiments explicitly examining target and distractor strength differentiation using a within-subjects design. She trained subjects in several CM conditions and then investigated the effects of target transfer (pairing trained CM targets with VM items) and distractor transfer (pairing VM items as targets with trained CM distractors). Positive transfer was demonstrated when either the CM target set or the CM distractor set remained the same and was paired with a VM set. These results demonstrated both target and distractor learning in visual search tasks.

Further evidence for both target and distractor learning in visual search has come from negative transfer (i.e., disruption in performance) found in studies that reversed the role of targets and distractors. Included in Dumais' (1979) experimental series were "partial reversal" conditions. A partial reversal is defined as a condition in which the role of either the target or the distractor set (but not both) has been reversed within a single condition. A target reversal involves using previously trained CM targets as distractors and pairing them with novel stimuli as targets. The CM stimuli, which have a high strength level, draw attention away from the new targets and serve to disrupt performance. Similar disruptions are found with distractor reversals in which the CM distractors become targets and are paired with novel items as distractors.

The strongest reversal effects, as would be expected from a strength perspective, were found in Shiffrin and Schneider's (1977, Experiment 1) "full reversal" condition. They trained CM targets and CM distractors and then reversed the roles of both the target and distractor sets within a single condition (i.e., previous CM targets became distractors for previous CM distractors, which then became the targets). Shiffrin and Schneider found that performance in the full reversal condition was actually worse than asymptotic VM performance. The large amount of disruption is consistent with the theory that attention is actually captured by the distractors and drawn away from the targets.

Another experiment in Dumais' (1979) series compared the differences in disruption due to full reversal and to partial reversals (i.e., target reversal and distractor reversal). Her results were consistent with Shiffrin and Schneider's in that full reversal yielded a strong disruption resulting in performance which was actually worse than asymptotic VM performance. She also found stronger disruption effects in the full reversal condition than in either of the partial reversals.

The experiments reviewed above provide supporting evidence that, within the visual search domain at least, subjects learn to attend to target information through strengthening or prioritizing that information. Furthermore, distractor information is ignored; hence, its attention-calling strength is reduced or weakened. These findings provide important information regarding the transfer of well-learned components to situations in which the utilization of the components remains similar (and performance is facilitated) or is reversed (and performance is disrupted). In a related manner, patterns of transfer and/or reversal allow estimation of the degree to which the components have been learned. This theoretical and empirical base was used in the present experimental series to investigate the effects of practice on the learning and transfer of components in visual search.

Outline of Experiments

The present experimental series was designed to investigate the effects of differential amounts of practice on the resultant strength of the CM items. In the first experiment, four groups of participants were trained for varying amounts of trials (560, 1120, 2240, or 3360 trials) in three semantic category visual search conditions--two CM and one VM. Following training, participants were tested in target, distractor, and full reversal conditions (cf. Rogers, 1989). As described above, the reversal of well-learned components results in disruption of performance. Thus, comparisons of the degree of disruption between the training groups allows a comparison of the degree of original learning.

Experiment 2 was an extension of the first experiment. A within-subjects, within-blocks design was used in which each participant received training in each of the following conditions: CM High (3150 trials), CM Moderate (1575 trials), CM Low (525 trials), and VM (1050 trials). Following training, there were two sessions of transfer which allowed a more complete specification of the effects of transfer and reversal of previously acquired automatic processes of varying strengths. The degree of disruption or transfer was measured as a function of different recombinations of items. For example, performance in six different target reversal conditions was measured to compare the amount of disruption in a target reversal situation in which the items used as distractors (i.e., previously trained CM targets) were manipulated. The distractors were either all highly trained CM targets, all moderately trained CM targets, all low trained CM targets or some combination of the three. Similarly, performance was measured for all combinations of distractor transfer.

Experiment 3 was similar to Experiment 2 except that training of the CM conditions was manipulated between blocks. We were interested in examining whether or not the relatively small differences between the CM High, CM Moderate, and CM Low conditions found in Experiment 2 were a function of the type of randomized training, training which may have allowed the development of a superset. In other words, it may have been possible for participants to create a superordinate category which contained all of the CM target categories. Thus, while the CM High category appeared most frequently as the target, the CM Moderate, and CM Low categories may have also been activated due to associative learning; thus, they would have benefitted from training to a greater degree than would be expected given the actual number of trials. This issue is explored in greater detail later in this report.

To summarize, the goal of the present series of experiments was to investigate the effects of the amount, type, and presentation (i.e., randomized or blocked) of practice on a visual search semantic category task. The focal points of interest were the effects on the development of automaticity, the benefits of additional training, and possible strategy differences among training situations.

Experiment 1 - Method

Subjects. Thirty-two subjects (14 males, 18 females) participated in the experiment. They received course credit for up to 6 hours of participation and/or were paid $4.00 per hour with a $1.00 per-hour bonus for completing the entire experiment. The vision of all participants was tested using a Snellen chart and their corrected or uncorrected visual acuity was at least 20/30 for distance and 20/40 for near (magazine print) vision.

Stimuli and Apparatus. Memory set items were the semantic unrelated categories (Collen, Wickens, & Daniele, 1975) of Animals, Vegetables, Units of Time, Countries, Body Parts, Weapons, Earth Forms, and Clothing. Target and distractor items were high associates of the categories (Battig & Montague, 1969). Each category consisted of eight words, each of which contained four to seven letters. Each participant received a unique assignment of categories for each condition; assignment was made using a Latin square counterbalancing design.

All stimuli were presented using EPSON Equity I+ microcomputers with Epson MBM 2095-5 green monochrome monitors. The standard Epson Q-203A keyboard was altered so that the '7' '4', and '1' numeric keypad keys were exchanged with the 'T', 'M', and 'B' keys, respectively. During all experimental sessions, pink noise was played at approximately 55 decibels (db) to help eliminate possibly distracting background noise. All participants were tested in the same room at individual, partitioned workstations monitored by a laboratory assistant.

Procedure. During the first session, participants were given practice consisting of five blocks of CM training (250 trials). These orientation trials allowed participants to become familiar with the experimental protocol and also served to stabilize error rates. The categories used for the orientation trials were not used in the remainder of the experiment.

An experimental trial consisted of the following sequence of events. A memory set item (i.e., category label) was presented for a maximum of 20 seconds, or until the participant pressed the space bar to initiate the remainder of the trial. Three vertically aligned plus signs were then presented for 0.5 second in the center of the screen to allow the participant to localize his/her gaze. The display set appeared in the same location as the plus signs and consisted of three category words presented in a column. The subject's task was to indicate the location of the target (top, middle, or bottom) by pressing the corresponding key labeled 'T', 'M', or 'B'. A target (i.e., an exemplar from the target category) was present on every trial.

Participants received the following performance feedback. After each correct trial, the response time (RT) was displayed in hundredths of a second. After each incorrect trial, an error tone sounded and the correct response was displayed. Following each block of 42 trials, the subject's average RT and percent accuracy for that block were displayed. If a subject's mean accuracy for a block fell below 90%, a warning message was displayed which encouraged him/her to respond more carefully. Participants were encouraged to maintain an accuracy rate of 95% while responding as quickly as possible. Participants were also encouraged to take short breaks between blocks. Before each session, participants were given feedback on the previous day's performance.

Design. There were two phases of the experiment: training and transfer. The three training conditions were: CM1 - A(B); CM2 - C(D); and VM - EFGH(EFGH) [where, for example, the representation A(B) refers to Target Set A displayed with Distractor Set B]. Within each session, participants completed seven blocks of each training condition. The order of presentation was as follows: CM1, CM2, VM. This sequence was repeated seven times, for a total of 21 blocks (42 trials per block).

The number of training sessions was manipulated between subjects. There were four groups: the 12-session group received 3360 trials per training condition, the 8-session group received 2240 trials per training condition, the 4-session group received 1120 trials per training condition, and the 2-session group received 560 trials per training condition. Eight participants were assigned to each of the four training groups.

After training, participants were placed in the transfer phase of the experiment. At the beginning of the transfer phase, participants were informed that the experimental conditions were going to change and that subsequent categories would appear in different pairings. They were also instructed to maintain their accuracy rates at 95% and to continue to try to respond as quickly as possible. The testing procedure used in the transfer phase of the experiment was identical to the procedure used in the training phase. The transfer session contained 20 blocks (5 per condition) of 42 trials each. Each participant completed 210 trials per condition, for a total of 840 transfer trials. The transfer conditions were: Full Reversal - B(A); Target Reversal - E(C); Distractor Reversal - D(F); and New CM - G(H).

In the reversal conditions, the roles of the targets and/or distractors were changed. In the full reversal condition both the CM target and CM distractor roles were reversed within a single condition [i.e., A(B) became B(A)] whereas in the partial reversal conditions (Target Reversal and Distractor Reversal) the role of <u>either</u> a CM target set <u>or</u> a CM distractor set was changed and each was paired with a previously trained VM set [e.g., in the target reversal condition, C(D) becomes E(C)]. In the New CM condition, two of the VM categories were paired to create a new consistently mapped condition to be used as a comparison condition. All transfer conditions were manipulated within subjects and presentation order of the conditions was randomized. The training conditions and the corresponding transfer conditions are summarized in Table 1.

The between-subjects independent variable was the amount of training provided: 12 sessions (3360 trials per condition), 8 sessions (2240 trials per condition), 4 sessions (1120 trials per condition), or 2 sessions (560 trials per condition). The within-subjects independent variables were: (a) Training condition (CM1, CM2, and VM); and (b) Transfer condition (Full Reversal, Target Reversal, Distractor Reversal, and New CM). The dependent variables were reaction time (RT) and accuracy.

<u>Experiment 1 - Results</u>

<u>Training Results</u>. Mean RT scores for each search condition for each training group as a function of practice are presented in Figure 1. An analysis of variance (ANOVA) for the first session of training yielded no significant differences among the four training groups. This was an important finding because it allows us to assume that all the training groups started at an equal level of performance. The main effect of Training

Table 1. Training and Transfer conditions

Training		Reversal
A (B)	--->	B(A) - Full Reversal
C (D)	--->	E(C) - Target Reversal
	--->	D(F) - Distractor Reversal
EFGH (EFGH)	--->	G(H) - New CM

NOTE: The representation A(B), for example, refers to Target Set A displayed with Distractor Set B.

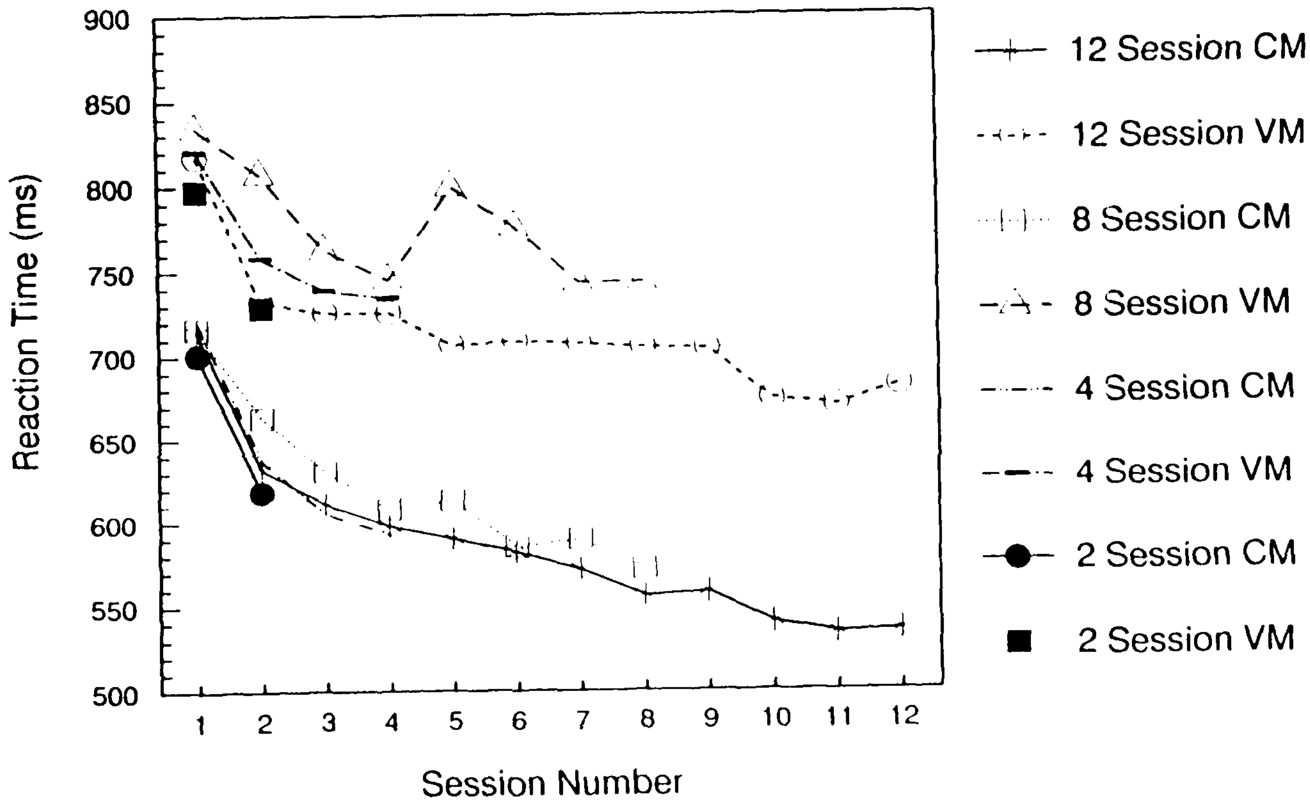


Figure 1. Reaction Time for Each Training Group in Each Search Condition, Plotted as a Function of the Number of Sessions.

condition was significant, $F(2,56) = 46.28$, $p < .0001$. The source of this effect was the fact that even after fewer than 300 trials, response times for the CM conditions were faster than those for the VM condition.

A Training condition (CM1, CM2, or VM) x Practice (Sessions 1 through 7) ANOVA was conducted for each of the four training groups (see Table 2). Each group showed significant performance improvements (as evidenced by the Practice main effects) as well as differences between CM and VM (as evidenced by significant Training Condition effects). Furthermore, for the training groups who received 12, 8, or 4 sessions of practice, the Training Condition x Practice interaction was significant. This interaction represented the differential rates of improvement for CM and VM (with CM, of course, improving to a greater degree). The fact that the 2-session training group did not show a significant Training Condition x Practice interaction was important, as it suggested that both the CM and VM conditions were improving at the same rate. This implied that more than 560 trials of training may be necessary before there is evidence of differential improvements between CM and VM. In other words, improvements early in training may be in a large part due to task-specific learning such as the location of the response keys, where to look on the screen, etc.

In order to measure the amount of improvement in performance due to training, we calculated the percent improvement for each participant ((First session RT - Last session RT) / First session RT) X 100. The aggregates of these functions are presented in

Table 2. Training Condition (CM1, CM2, or VM) x Practice (Sessions 1 through 7) Analysis of Variance

Training group	Training condition	Practice	Training condition x Practice
12-session	$F(2,14)=13.51^*$	$F(11,77)=36.47^*$	$F(22,154)=2.67^*$
8-session	$F(2,14)=27.25^*$	$F(7,49)=17.71^*$	$F(14,98)=3.73^*$
4-session	$F(2,14)=19.27^*$	$F(3,21)=66.51^*$	$F(6,42)=3.28^*$
2-session	$F(2,14)=13.78^*$	$F(1,7)=13.78^*$	Not significant

$^*p<.01$

Table 3. The average accuracy rate for the first session was 96% (range 94%-97%) and for the final session, 95% (range 93%-96%). For the purposes of this discussion, we will focus mainly on RT scores. These data were commensurate with the ANOVA results reported above: Namely, there was greater improvement in CM than in VM for the 12-session, 8-session, and 4-session groups. Furthermore, averaged across the two CM conditions, Student-Newman-Keuls comparisons showed that the 12-session group improved the most (25%), followed by the 8-session and 4-session groups which did not differ significantly (20% and 18%, respectively); the 2-session group improved the least (12%).

To summarize the training results, we can focus on several emergent patterns. First, as expected, CM practice produces generally faster performance than VM practice. However, given the fact that the 2-session training group showed no Practice x Training condition interaction, differential improvement rates for CM relative to VM may not be evident very early in training. This is not surprising given that, early in training, both CM and VM tasks are dominated by controlled processing. The second general pattern of training results demonstrates that more practice is beneficial, in that the 4-, 8-, and 12-session groups all showed performance superior to that of the 2-session group; however, after practice, performance in the 4-, 8-, and 12-session groups was not significantly different. The possibility that 4 sessions of training provide benefits equivalent to those produced by 12 sessions is explored in the following section.

Table 3. Improvement in RT with Practice (Difference between first and last sessions of training)

	Training condition		
Training group	CM1	CM2	VM
12-session	26%	24%	17%
8-session	21%	20%	11%
4-session	20%	16%	10%
2-session	8%	11%	9%

Transfer Results. The main point of interest here was the effects on performance of reversing either the targets or the distractors or both. We measured the degree of disruption in two ways. First, within each training group, we compared performance in the Target Reversal, Distractor Reversal, and Full Reversal conditions to the New CM control condition, using planned comparisons. Typically previous research has demonstrated that the target reversal and distractor reversal conditions yield performance which is equivalent a new CM condition, thereby implying a reversion to controlled processing. In the present experiment all four training groups yielded this pattern; that is, the differences between Target Reversal, Distractor Reversal, and New CM were not significant. Research has demonstrated that the largest reversal effects occur in the Full Reversal condition such that performance is worse than that for New CM (cf. Shiffrin & Schneider, 1977). In the present experiment, the comparisons of Full Reversal to New CM were significant for the 12-session and 8-session training groups, $F(1,21) = 20.69$, $p < .0002$ and $F(1,21) = 8.91$, $p < .007$, respectively. For the 4-session group, the difference between Full Reversal and New CM was marginally significant, $F(1,21) = 3.69$, $p < .068$. The contrast was not significant for the 2-session group ($F < 1$).

As pointed out earlier, the amount of disruption in reversal conditions is an indicator of the degree of original learning. The present pattern of reversal effects supports earlier claims that, to some degree, more practice is better. The 2-session group showed the least amount of disruption, which was indicative of the least amount of original learning. The 12-session and 8-session groups displayed the strongest effects and the 4-session group demonstrated a slightly weaker effect.

A second measure of the amount of performance disruption due to reversal was obtained by calculating the difference between final CM training performance and each of the reversal scores (i.e., {[CM1-Target Reversal]/ CM1} * 100 yields the percentage of change in performance due to Target Reversal). These data are presented in Table 4. The pattern of disruption corresponded to the results of the RT analysis reported above. The Full Reversal condition showed the most disruption and the amount of disruption decreased as the amount of original training decreased.

<u>Experiment 1 - Discussion</u>

In general, the results of the present experiment suggest that <u>more</u> practice is better. However, several caveats must accompany this statement. First of all, more practice is better only if the role of the trained components will not be reversed at some point following training. As we demonstrated, well-learned components do yield superior performance; however, they are also more difficult to "unlearn" if necessary (see Section IV of this report). A second caveat is related to the question of how much more is better. Clearly, the participants in the 2-session group did not show much benefit from the amount of training they received (see Table 3). In fact, for this group the mean improvement for the CM conditions (10%) was hardly better than the general improvement for the VM training condition (9%). It is also evident that although the 12-session group did

Table 4. Percent of Disruption in RT for Each Reversal Condition Relative to the CM Training Condition (CM RT-Reversal RT / CM RT)

	Reversal condition		
Training group	Target Rev.	Distractor Rev.	Full Rev.
12-sessions	20%	30%	46%
8-sessions	22%	28%	36%
4-sessions	20%	19%	25%
2-sessions	6%	9%	13%

show significant improvement (as evidenced by both improvement functions and reversal patterns), it is not clear that they benefitted much more than did the 8-session group.

Finally, there is another issue which cannot be directly addressed by the present experiment. This relates to the potential benefit of providing general training on related tasks for producing improvement in overall performance. The between-subjects design of the present experiment, in effect, confounded amount of time on task with practice per condition. In Experiment 2, we addressed the issue of component-specific training versus general training time.

Randomized vs. Blocked Practice. Recently, Carlson, Sullivan, and Schneider (1989) reported the results of an experiment investigating the acquisition of skill in making logic gate judgments based on a series of rules. Participants were initially provided with blocked practice in which rules were learned one at a time (see also Carlson & Yuare, 1988). Participants then practiced with the same rules mixed together within blocks of trials. There was a large decrement in performance in the transition from blocked practice on each rule (i.e., logic gate) to practice in which the rules were randomized within blocks. Carlson et al. (1989; and Anderson, 1989, in his comment on the paper) suggested that in the blocked practice, participants need only to establish associations between possible input patterns and output values without first identifying the gate type (i.e., they can use a single-step associative process), but that in randomized practice, the participants may have to use a serial judgment process due to the need to first identify the gate type.

The Carlson et al. (1989) results suggest that perhaps the benefits of blocked practice may not carry over to a situation which involves the randomization of trials. Related to this point is the question of whether between-block manipulations allow (or require) different strategies of performance than within-block manipulations. A slight variation of this issue was investigated in the following two experiments. The chief manipulation was the amount of practice provided for each CM condition. An additional manipulation involved whether all of the conditions were combined in a block of trials or instead were presented in pure unmixed blocks. To illustrate the distinction, in Experiment 2 each block contained 50 trials: 30 trials of the CM High condition, 15 trials of the CM Moderate condition, and 5 trials of the CM Low condition (the trials were randomly intermixed). In this situation, it might be possible for the participants to create a superordinate category which contains all of the CM target categories. Thus, though the CM High category would appear most frequently as the target, the CM Moderate and CM Low categories might also be activated due to associative learning. In Experiment 3, the training conditions were presented in separate blocks of 40 trials. Thus an entire block consisting of the CM High category might be followed by a block of the CM Moderate category and then by a block of the CM Low category. (The actual order of presentation was random.) In order to provide the requisite number of training trials for each of the conditions, the number of blocks was manipulated. In this situation, the opportunity to create a superordinate category is not present.

Also manipulated in Experiments 2 and 3 were the effects of different recombinations of components on performance. That is, there were two transfer sessions in which previously trained components were paired either in target reversal conditions or in distractor transfer conditions. For example, one target reversal condition might consist of a former highly trained CM category as one distractor, as well as a former moderately trained CM category as another distractor. All such combinations were included in an effort to specify more precisely the effects of differential training of task components in various transfer situations.

Experiment 2 - Method

Subjects. Fourteen new subjects (7 male, 7 females) participated in the experiment. The participants were compensated monetarily for their participation: $4.00 per hour, with a $1.00-per-hour bonus for completing the entire experiment. The vision of all participants was tested using a Snellen chart and their corrected or uncorrected visual acuity was at least 20/30 for distance and 20/40 for near (magazine print) vision.

Stimuli. Memory set items were the semantically unrelated categories (Collen et al., 1975) of Furniture, Vegetables, Musical Instruments, Four-Footed Animals, Alcoholic Beverages, Building Parts, Weapons, Earth Formations, Units of Time, Occupations, Body Parts, Relatives, Vehicles, Countries, Trees, and Clothing. Target and distractor items were high associates of these categories (Battig & Montague, 1969). Each category set contained eight words. Each participant received a unique assignment of categories for each condition, counterbalanced by a Latin square.

Apparatus. The apparatus was identical to that of Experiment 1.

Procedure. During the first session of the experiment, the participants completed a practice session of the experimental task. The practice session consisted of five blocks of CM trials (50 trials per block). These orientation trials allowed the participants to become familiar with the experimental protocol and also served to stabilize the error rates. The categories used for the practice trials were not used in the remainder of the experiment.

An individual trial consisted of the following sequence of events. The participant was presented with the memory set of one category label, which he/she was allowed to study for a maximum of 20 seconds. Participants were instructed to press the space bar to initiate the trial. Three plus signs were then presented in a column for 0.5 second in the location of the display set (in the center of the screen) to allow the participant to localize his/her gaze. The plus signs were followed by the display set, which consisted of three words presented in a column. The subject's task was to indicate the location of the target (i.e., top, middle, or bottom) by pressing the corresponding key (labeled 'T', 'M', or 'B'). A target (i.e., an exemplar from the target category) was present on every trial.

Participants received the following performance feedback. After each correct trial, the subject's RT was displayed in hundredths of a second. After each incorrect trial, an error tone sounded and the correct response was displayed. Following each block of trials, the participant received his/her average RT and percent accuracy for that block; if a subject's accuracy fell below 90% in any block, a message was displayed encouraging him/her to respond more carefully. Participants were instructed to maintain an accuracy rate of 95% or better while responding as quickly as possible. After each block of trials, participants were encouraged to take a short break to rest their eyes.

There were two phases of the experiment: training and testing. The training phase consisted of four conditions: (a) CM High - 3150 trials, (b) CM Moderate - 1575 trials, (c) CM Low - 525 trials, and (d) VM - 1050 trials. The participants were trained for seven 1-hour sessions, each of which consisted of 15 blocks of CM training (50 trials per block - 30, 15, and 5 trials for each of the CM conditions, respectively, which were presented in a random order) and five blocks of VM training (30 trials per block). Three CM blocks were presented, followed by one block of VM; this sequence was then repeated four more times to complete a session.

The testing phase of the experiment consisted of two sessions: one session of Target Reversal conditions and one session of Distractor Transfer conditions. In the Target

Reversal (TR) conditions, previously trained VM sets were used as target items and the types of distractors (i.e., previously CM High, Moderate, or Low trained target items) were manipulated. The conditions were as follows:

1. High/High Target Reversal - both distractor items on a trial were previously CM High targets.
2. Moderate/Moderate Target Reversal - both distractor items on a trial were previously CM Moderate targets.
3. Low/Low Target Reversal - both distractor items on a trial were previously CM Low targets.
4. High/Moderate Target Reversal - one distractor item was previously a CM High target and the other was previously a CM Moderate target.
5. High/Low Target Reversal - one distractor item was previously a CM High target and the other was previously a CM Low target.
6. Moderate/Low Target Reversal - one distractor item was previously a CM Moderate target and the other was previously a CM Low target.
7. New CM condition - created by pairing two of the VM sets in a consistent mapping.

The New CM condition served as a comparison condition. The six target reversal conditions were manipulated within a block and the New CM condition was presented in a separate block. Three blocks of Target Reversal were completed, followed by one block of the New CM condition; a single session consisted of five repetitions of this sequence. Participants completed 15 blocks

(60 trials per block), for a total of 150 trials for each of the six target reversal conditions and 5 blocks (30 trials per block) for the New CM condition.

In the Distractor Transfer (DT) conditions, previously trained VM sets were used as target items and the types of distractors (i.e., previously CM High, Moderate, or Low trained distractor items) were manipulated. The conditions were as follows:

1. High/High Distractor Transfer - both distractor items on a trial were previously CM High distractors.
2. Moderate/Moderate Distractor Transfer - both distractor items on a trial were previously CM Moderate distractors.
3. Low/Low Distractor Transfer - both distractor items on a trial were previously CM Low distractors.
4. High/Moderate Distractor Transfer - one distractor item was previously a CM High distractor item and the other was previously a CM Moderate distractor.
5. High/Low Distractor Transfer - one distractor item was previously a CM High distractor item and the other was previously a CM Low distractor.
6. Moderate/Low Distractor Transfer - one distractor item was previously a CM Moderate distractor item and the other was previously a CM Low distractor.
7. New CM condition - created by pairing two of the VM sets in a consistent mapping.

The New CM condition was included as a comparison condition. The six Distractor Reversal conditions were manipulated within a block and the New CM condition was presented in a separate block. Three blocks of Distractor Reversal were completed, followed by one block of the New CM condition; a single session consisted of five repetitions of this sequence. Participants completed 15 blocks (60 trials per block), for a total of 150 trials per Distractor Reversal condition and 5 blocks (30 trials per block) of the New CM.

Design. Within-subjects independent variables were: (a) Training conditions: CM High, CM Moderate, CM Low, and VM; (b) Target Reversal conditions: High/High Target Reversal, Moderate/Moderate Target Reversal, Low/Low Target Reversal, High/Moderate Target Reversal, High/Low Target Reversal, Moderate/Low Target Reversal, and New CM; and (c) Distractor Transfer conditions: High/High Distractor Transfer, Moderate/Moderate Distractor Transfer, Low/Low Distractor Transfer, High/Moderate Distractor Transfer, High/Low Distractor Transfer, Moderate/Low Distractor Transfer, and New CM. The CM, Target Reversal, and Distractor Transfer conditions were manipulated within blocks whereas VM and New CM were manipulated between blocks. The dependent variables were RT and accuracy.

Experiment 2 - Results

Training Results. An ANOVA was performed on the RT scores for the first session of training. There was a significant effect of Training condition, $F(3,39) = 6.42$, $p < .0012$. Contrasts between the Training conditions revealed that the CM

High, CM Moderate, and CM Low condition were all significantly different from VM, $\underline{F}(1,39) = 18.38$, $\underline{p} < .001$, but they did not differ significantly from each other.

An ANOVA was also computed for the last session of training. Again, there was a significant effect of Training condition, $\underline{F}(3,39) = 25$, $\underline{p} < .001$. However, the source of this effect was not entirely due to differences between the CM and VM conditions. A Student-Newman-Keuls test of comparisons at the .05 level revealed that although all of the CM conditions were significantly different from the VM condition, both the CM High and CM Moderate conditions resulted in significantly faster performance than the CM Low condition, but were not significantly different from each other.

The fact that the CM conditions were equivalent during the first session but significantly different after training suggests that the CM training conditions improved differentially (see Table 5 for the mean percentages of improvement). This was confirmed by the Training condition x Practice ANOVA. As expected, when the analysis was run with all of the training conditions included there were significant main effects of Training condition, $\underline{F}(3,39) = 21.13$, $\underline{p} < .0001$, and Practice, $\underline{F}(6,78) = 50.81$, $\underline{p} < .0001$, as well as a significant Training condition x Practice interaction, $\underline{F}(18,234) = 3.56$, $\underline{p} < .0001$. It is noteworthy that the Training condition x Practice interaction remained significant even when the analysis was run without the VM condition, $\underline{F}(12,156) = 2.25$, $\underline{p} < .01$. This

Table 5. Improvement in RT with Practice

	Training condition			
	CM High	CM Moderate	CM Low	VM
Beginning RT	744	740	755	810
Ending RT	568	585	614	700
% Change	24%	21%	19%	14%

supports the contention that there were differential rates of improvement among the CM training conditions.

A Training condition x Practice ANOVA on the accuracy data yielded significant main effects of Training condition, $F(6,78) = 2.47$, $p < .03$, and Practice, $F(3,39) = 7.79$, $p < .0003$, but the interaction was not significant. The average accuracy for the CM conditions was 95%, which was slightly better than the VM condition (93%). Furthermore, there was a slight decrease in accuracy across sessions from 95% to 94%.

Target Reversal. An ANOVA conducted on the RT data yielded a significant effect of Transfer condition, $F(6,78) = 3.78$, $p < .003$. Further probing with a Student-Newman-Keuls test showed that all of the reversal conditions were significantly slower than the New CM condition but not significantly different from each other. Thus, regardless of the pairings of the items, if former CM targets (whether High, Moderate or Low trained) were used as distractors, they were disruptive to performance. In other words, the participants were unable to ignore the previously attended items. The accuracy scores ranged from 92% to 95%, but there were no clearly meaningful patterns of differences among the conditions.

Distractor Transfer. An analysis of variance of the RT data for the Distractor Transfer session also yielded a significant effect of Transfer condition, $F(6,78) = 3.13$, $p < .0084$. The Student-Newman-Keuls test of these data revealed that all of the transfer conditions showed significantly faster performance than the New CM condition but were not significantly different from each other. The accuracy scores ranged from 92% to 94%, and there were no significant differences among the conditions. These results suggested that participants did benefit from having ignored the distractor items previously, and that they were able to transfer to new target items while maintaining their level of performance.

Experiment 2 - Discussion

The training results are typical in that response times in the CM conditions were faster than those in the VM condition. This was true even for the CM Low condition, which received only 525 trials. Furthermore, the results showed that the CM High (3150 trials) and CM Moderate (1575 trials) conditions were relatively equivalent, but they were both better than the CM Low condition. Thus, it is again evident that at least to some extent, more practice results in better, or at least faster, performance.

However, the patterns of Target Reversal and Distractor Transfer tell a somewhat different story. The Target Reversal conditions, which involved the reversal of CM Low items, yielded an amount of disruption equivalent to those involving the reversal of CM High items. This pattern would not be expected if in fact the CM High items were learned to a greater degree. It was possible in this experiment, however, for participants to form a superset of the CM categories because they were trained in a within-block design. Thus, the CM Low items may have been activated to a greater degree simply due to the associative learning that generally takes place during CM training. This possibility was explored in the following experiment in which the opportunity to form a superset was removed (i.e., the conditions were trained in separate blocks).

Experiment 3 - Method

Subjects. Seven new subjects (5 males, 2 females) participated in the third experiment. Participants received course credit for up to 6 hours of participation and/or $4.00 per hour, with a $1.00 per-hour bonus upon completion of the entire experiment. The vision of all participants was tested using a Snellen chart, and their corrected or uncorrected visual acuity was at least 20/30 for distance and 20/40 for near (magazine print) vision.

Stimuli and Apparatus. The stimuli and apparatus were identical to those of Experiment 2.

Procedure. The procedure for individual trial presentation and feedback was identical to that of Experiment 2. The major difference from Experiment 2 was the manipulation of training conditions; in Experiment 2, the conditions were manipulated within blocks whereas conditions in the present experiment were manipulated between blocks. There were two phases of the experiment: training and testing. The training phase consisted of four conditions: (a) CM High - 3360 trials, (b) CM Moderate - 1,680 trials, (c) CM Low - 560 trials, and (d) VM - 1,120 trials. The participants were trained for seven 1-hour sessions, each of which consisted of 24 blocks (40 trials per block): 12 blocks of CM High, 6 blocks of CM Moderate, 2 blocks of CM Low, and 4 blocks of VM. The order of presentation of the blocks was randomized.

The testing phase of the experiment was the same as that used in Experiment 2, with the following exception. Four blocks (48 trials) of each transfer condition (target reversals in the first transfer session and distractor transfers in the second transfer session) were presented, followed by a block of the New CM condition (32 trials). Five repetitions of the sequence made up each of the transfer sessions.

Experiment 3 - Results

Training Results. An ANOVA was performed on the RT scores for the first session of training. There was a significant effect of Training condition, $F(3,18) = 4.42$, $p < .017$. The Student-Newman-Keuls comparisons revealed that the CM High, CM Moderate, and CM Low conditions were all significantly different from VM but did not differ significantly from each other.

An ANOVA was also performed on data from the final session of training (see Table 6). Again, there was a significant effect of Training condition, $F(3,39) = 25$, $p < .001$. The source of this effect was entirely due to differences between the CM and VM conditions.

General Discussion - Experimental Series 1

The present series of experiments was designed to answer a critical question for training programs: How much practice is necessary and in what form should it be in order to maximize learning? Not surprisingly, we found that, generally speaking, more practice is better. However, several caveats must accompany

Table 6. Improvement in RT with Practice

	Training condition			
	CM High	CM Moderate	CM Low	VM
Beginning RT	732	779	739	878
Ending RT	576	583	610	696
% Change	21%	25%	18%	21%

such a statement. First of all, in the process of using reversal conditions to measure learning, we discovered that the pattern of disruption effects also demonstrated that overtraining components of a task will disrupt performance if the role of these components changes (cf. Dumais, 1979; Fisk & Rogers, 1988; Lee, Rogers, & Fisk, in press).

A second finding evident in the present data was that although 1,000 trials were better than 500 trials, and 3,000 trials were better than 2,000 trials, 2,000 trials were not necessarily better than 1,000 trials. These apparently discrepant findings may be reconciled if one views the transition from novice to skilled performance or controlled to automatic processing as passing through several stages. Fitts and Posner (1967), for example, referred to this transition as moving from cognitive to associative to autonomous information processing. Similarly, Anderson (1982, 1983), in his production system model, described the stages of transition as the declarative stage, the knowledge compilation stage, and the procedural stage. Recently, Schneider and Detweiler (1987) further specified the transitions of controlled to automatic processing. They described five phases: Phase 1 - controlled comparisons from buffered memory; Phase 2 - context-maintained controlled comparison; Phase 3 - goal-state-maintained control comparison; Phase 4 - controlled assist of automatic processing; and Phase 5 - pure automatic processing.

It is conceivable, therefore, that differences in performance after differing amounts of practice simply imply that the individual is in a stage of transition from controlled to automatic processing. Thus, for example, in Experiment 1, the participants in the 4-session and 8-session training groups may have been performing the task at adjacent stages of processing and thus differences between them were not detected in the simple RT comparisons. However, in terms of reversal effects, the 8-session group did show more severe disruption, thus signifying somewhat greater improvement for that group.

The purpose of discussing the transition from controlled to automatic processing is to illustrate that more consistent practice may be a key contributor to the progression from one stage of information processing to another. In light of this, the results of Experiment 2 might seem surprising. Why was there no significant difference between performance in the CM High (3,150 trials) and CM Moderate (1,575) training groups? In that experiment, the training for each of the conditions was randomized within blocks. We have speculated that this form of practice allows the Low and Moderate conditions to benefit from the frequent occurrence of the High category due to associative learning. In CM practice situations, memory-set items are associatively connected to form a superset. After practice, the activation of one member of the set results in the associative activation of the other members, thus strengthening them (i.e., increasing their priority level).

The possibility that supersets may be formed during randomized within-block presentation of conditions has exciting implications for training programs. Suppose, for example, that one wishes to train participants to detect a series of movement patterns or to learn groups of symbols which have no apparent relationship (recall that in the present experiments the categories were <u>not</u> related). By providing within-block training, one can capitalize on the fact that associative learning between memory-set items generally takes place early in CM training. Thus, a large amount of training need not be required for each and every stimulus but only for the superset as a whole. Associative learning may allow all members of a superset to benefit from practice on any of the members. For pure visual search, it is likely that some base amount of practice will be necessary on each of the exemplars for the purposes of feature differentiation and identification.

This is one of several avenues of future investigation in this area. Other topics of interest will be a more precise specification of the time course of associative learning and priority learning, and designation of the processes of associative activation. Whether or not associative activation will be sufficient to increase or decrease the priority tag of a stimulus or class of stimuli is open to investigation. More than likely, associative activation will be necessary but not sufficient and some level of direct activation will be required. This provides a third avenue for exploration. The present results indicate that perhaps the overall amount of practice necessary to reap maximum benefits might be reduced by "packing" training and allowing associative learning to spread across the members of a superset which may be defined by the trainer.

III. EXPERIMENTAL SERIES 2: TRANSFER OF TRAINING AS A FUNCTION OF SEMANTIC RELATEDNESS IN A CATEGORY SEARCH TASK

Introduction

For the better part of this century, transfer of training was one of the most heavily studied phenomena within psychology (e.g., Briggs, 1969; Bruce, 1933; Osgood, 1949; Thorndike & Woodworth, 1901a, 1901b, 1901c). With the ascent of cognitive psychology, interest in this important topic waned; recently, however, interest in transfer of training issues seems to have made a comeback (e.g., Cormier, 1987; Gick & Holyoak, 1987; Gray & Orasanu, 1987; Schneider & Fisk, 1984).

A key issue in the area of training involves determining how, after extensive training in performance of a task, people perform when faced with a novel, but related task. The present investigation examined this issue by testing the transfer of highly trained, automatized components of a semantic category search task to components of varying degrees of relatedness. From a training perspective, this issue is critical to many real-world "high performance" skills (Schneider, 1985a). Consider, for example, learning symbology, radio transmission calls or tactical formations from the perspective of an air weapons controller. Clearly, what is learned in one situation during training (particularly in radio transmission and tactical formations) may not repeat itself exactly in real missions. However, the class or category of events learned by the air weapons controller is generally consistent (or at least related to some degree) across situations.

In order to examine how training on high performance skills is transferred to novel but related tasks, an analysis of which performance characteristics change as skill develops is needed. As a result of investigations based on automaticity theory by Fisk and his colleagues, a picture of these changes is emerging. These contributions of automaticity theory to the areas of training and skill acquisition are well documented (e.g., Fisk & Eboch, 1987; Fisk & Lloyd, 1988; Myers & Fisk, 1987; Schneider, 1985a).

Performance characterized as "expert" or "skilled" develops only under CM conditions. However, in a real-world setting, complex tasks demanding high levels of skill are dependent on both controlled and automatic processing. Understanding the development and the maintenance of skilled performance requires analysis of the task in terms of the component processes that drive performance. Such an analysis is essential to the identification of those component processes which are transferrable to different but related tasks.

It is known that automatic processing can transfer across some situations. Within-category transfer of training effects, for example, have been demonstrated in memory search tasks (Schneider & Fisk, 1984). There is substantial evidence that most characteristics of automatic processes are well described as memory phenomena (Fisk & Rogers, 1988; Logan, 1988a, 1988b; Schneider & Fisk, 1984). Also, it is clear that high performance skill learning need not take place at the specific individual stimulus level (e.g., Fisk & Lloyd, 1988; Fisk, Oransky, &

Skedsvold, 1988; Myers & Fisk, 1987). However, the degree to which the characteristics of highly trained, automatic processes seen in memory search will be exhibited in visual search tasks has not been well established. Therefore, the present series of experiments was conducted to test the limits of transfer effects and "levels of learning" for various task requirements in visual search. Understanding performance and transfer effects in visual search tasks is important because many real-world, high performance skills rely heavily on a visual search component.

At a theoretical level, this investigation was designed to extend understanding of automatic processing by examining transfer of training on a semantic category search task. At a practical level, the results of this study will contribute data which instructional system designers could use in developing training programs.

Experiment 1 - Method

Subjects. Six right-handed psychology graduate students (5 males, and 1 female) from the Georgia Institute of Technology volunteered to participate in the study. Participants were tested for visual acuity of at least 20/30 (uncorrected or corrected) and near vision of at least 20/40. Participants were paid for their time.

Equipment. Epson Equity I+ microcomputers equipped with Epson MBM-2095-E monochrome monitors (green phosphor, 50-Hz refresh rate) and Epson multimode graphics adapters were used to present the task. The microcomputers were programmed with Psychological Software Tools' Microcomputer Experimental Language

(MEL) to present and time stimulus displays and to record response behavior.

There were five experimental stations. Each station consisted of three sound-deadening panels which formed a booth. A desktop on which a single microcomputer was placed was located within this booth. Each booth contained two speakers through which pink noise was played at a sound level of approximately 55 dB. Therefore, sounds external to an individual's booth were masked and participants could not see each other.

Stimuli. All categories and exemplars were from Battig and Montague's (1969) semantic category norms and were chosen according to the following criteria: (a) degree of semantic relatedness among categories (as determined by Collen et al., 1975), (b) exemplar length between four and seven letters, and (c) target exemplars of high to moderately high production frequency (high item dominance) ranking (Battig & Montague, 1969).

During training, participants searched for target words (eight exemplars from a single category) against a background of distractor words (exemplars from six categories semantically unrelated to the target categories). During transfer, four new target categories were presented (six exemplars per category), as well as six new exemplars from the category on which participants trained. Also, to avoid confounding of distractor learning (Dumais, 1979; Kristofferson, 1977; Rogers, 1989), 48 exemplars from six new distractor categories were used during the transfer sessions. (Please see Appendix A for a complete listing of categories, along with their exemplars, and Appendix B for the percentage of relationship among target categories.)

All words were presented in uppercase letters. Participants were seated approximately 48 centimeters (cm) from the display. At that viewing distance, the average letter subtended 0.38 degree in width and 0.47 degree in height. Within a word, interletter separation was 0.19 degree.

Procedure. In order to test performance at the limits of each individual's visual search capacity, we developed an adaptive version of the "multiple-frame" detection task for the training phase of this experiment. This task was based upon multiple-frame tasks used typically in the visual search/detection literature (e.g., Schneider & Shiffrin, 1977; Sperling, Budiansky, Spivak, & Johnson, 1971). However, in our version of the task, frame time (the time from the onset of one display until the onset of the next display) is determined by each individual's visual search accuracy. As accuracy changes, so will frame time. Although conceptually simple, the task is quite demanding. The procedure avoids some of the controversies associated with reaction time studies (e.g., speed-accuracy trade-offs).

All participants began the experiment at the same "speed," with frame time equal to 700 milliseconds (ms). From that point until the final block of the final session of training, each individual's performance (as measured by accuracy criteria) determined the frame speed for subsequent blocks. If a participant's accuracy on any block was equal to or better than

80% correct (24 correct out of a total of 30 trials), frame time on the next block was <u>decreased</u> by 25 ms. If accuracy fell below 80% (i.e., below 23 correct responses), frame time on the next block was <u>increased</u> by 25 ms. Accuracy on the final block of any training session determined the initial frame time on the first block of the subsequent training session. During transfer the frame time was held constant across all blocks. The frame time was determined by using the frame time from the last block of the final training session on which the accuracy criterion was attained (i.e., at least 80% correct). Thus, frame time was held constant during transfer and accuracy was the dependent measure.

A representation of a single, multiple-frame trial is provided in Figure 2. At the beginning of each trial, participants studied a memory set (a single semantic category) for a maximum of 20 seconds. Once the individual encoded the set, he or she initiated presentation of the frames by pressing the space bar. As the name implies, the frame was the main element of this procedure. A frame consisted of two displays presented sequentially. The first display consisted of the display set, which contained three semantic category exemplars displayed in a column through which the participant was to search. The second display contained a visual mask consisting of three rows of X's to prevent continued processing of the display set after its removal from the video display unit (VDU).

In this study, eight frames per trial were used. Each sequence of frames was presented following a 500-ms display of focus points (three plus signs (+) displayed in a column where

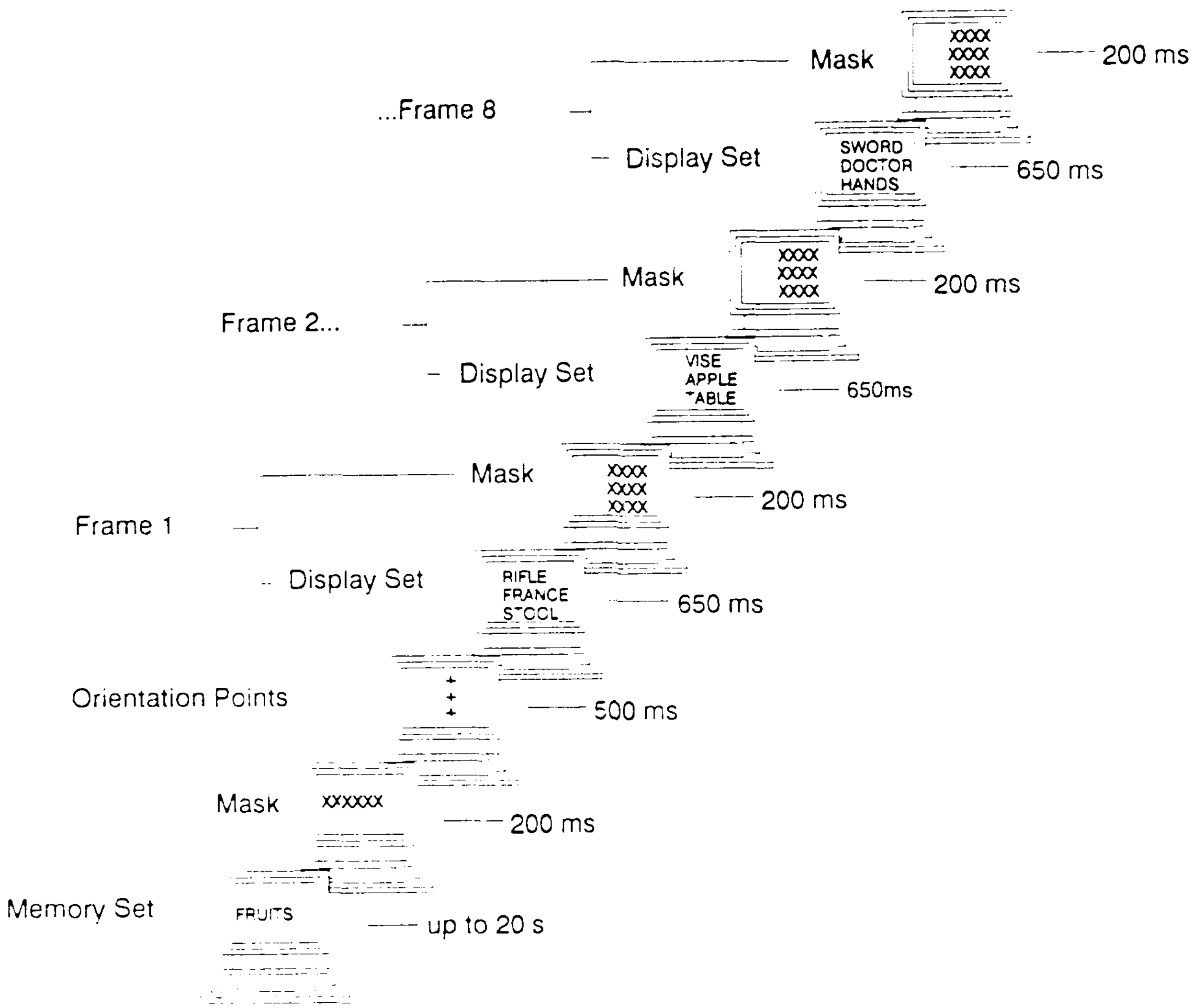


Figure 2. The Multiple-Frame Procedure. In this representation, Frames 3 through 7 are omitted. The target, "APPLE", appears in the middle position, on Frame 2. Hence, the correct response would be to press the key labeled "M".

the exemplars were displayed). Frame time was measured from the onset of display of one frame to the onset of the next frame (a zero interframe interval). Although presentation time for the display set varied across blocks as a function of an individual's accuracy, presentation time of the visual mask remained constant at 200 ms. The eight frames were displayed sequentially and rapidly, much like a slide projector with a stuck button (see Figure 2).

Participants searched through 24 exemplars (eight frames x three exemplars per frame) to find a target. There were two kinds of trials: target present (positive trials) and target absent (negative trials). On positive trials there was one target (an exemplar from the category which appeared in the memory set) appearing within Frames 2, 3, 4, 5, 6, or 7 (never in Frame 1 or Frame 8) in either the top, middle or bottom position on the VDU screen. Both frame number and vertical position were selected randomly. If the trial was positive, the correct response was to press a key labeled 'T', 'M' or 'B' (corresponding to the 7, 4 or 1 key on the numeric keypad) depending on the vertical location of the target exemplar. If the trial was negative, the correct response was to press a key labeled 'N' (corresponding to the 5 key on the numeric keypad).

Participants could respond at any point during presentation of the frames and for up to 4 seconds after the final frame. Following the response, the VDU screen was cleared and feedback for that trial was presented. After each trial, participants received correlated visual and auditory feedback about their response. If a correct response was entered, the microcomputer displayed the word "CORRECT!" inside a box at the left center of the screen. If the participant "missed," then the message "ERROR, exemplar was presented in position" (where exemplar was the actual target word and position was the actual vertical position of the target for that trial) was displayed at the target location, simultaneously with presentation of a 1200-Hz tone. If the participant "false-alarmed," then the microcomputer displayed "ERROR, there was no target present" in the right center of the screen, simultaneously with presentation of a 100-Hz tone. If the participant made an "error of position," then the microcomputer displayed "ERROR, exemplar was present in position" at the target location, simultaneously with presentation of a 500-Hz tone.

At the end of each block, participants received feedback and had an opportunity to take a break (and were encouraged to do so). First, information about performance on the just-completed block was displayed for 7 seconds. Then, cumulative feedback in the form of textual information about an individual's performance was displayed. When a participant was finished viewing the feedback screen, he or she pressed the space bar to initiate the next block of trials.

Design. It is important to note that the training phase was of minor significance; it is for the transfer phase where the issues of interest will be examined. As previously mentioned, participants were pushed to perform at their perceptual processing limits. Toward this end, much of the data collected served to provide us with daily reports on participants' progress during training to ensure that participants were indeed "keeping on task."

All manipulations were within-subject.[1] Data from the following indep ndent variables were collected: (a) position of the target (top, middle, bottom or no target present), (b) frame number of the target exemplar (two through seven), (c) type of trial (positive or negative), (d) target category (i.e., memory set), and (e) target exemplar.

The primary dependent variable during training was display set time ("speed") and during transfer, correct-incorrect response (accuracy). Also, the time spent encoding the memory set was collected.

This study was divided into two phases: training and transfer. Training consisted six training sessions. During the first training session, we obtained demographic and health information, tested visual acuity and instructed participants on how to perform the task. Training sessions consisted of 14 blocks of trials, with 30 trials per block. Twenty percent of all trials were negative (target absent). There were 420 trials per session, for a total of 2,520 trials (2,016 positive trials and 504 negative).

There was one transfer session consisting of six conditions manipulated across 12 blocks. The Priming condition consisted of

[1] Strictly speaking, this is not true: During training, half of the participants had "Fruits" as their target category and half had "Vegetables." Consequently, participants who trained on Fruits had Vegetables for their Highly Related (HR) condition and vice versa.

the same category and exemplars on which an individual had trained previously and was always presented in Blocks 1 and 2. The remaining five conditions were presented pseudo-randomly within Blocks 3 through 7, and again within Blocks 8 through 12, with the proviso that the same condition could not appear back-to-back. Conditions Trained/Trained (T/T) and Trained/Untrained (T/U) were presented within the same blocks. T/T consisted of the previously trained category exemplars (8 words from the training phase) and T/U consisted of the previously trained category but with untrained exemplars (6 new words). The remaining conditions consisted of both untrained categories and exemplars and were manipulated across blocks. The Highly Related (HR) condition contained six exemplars from a category that was highly semantically related (Collen et al., 1975) to the category on which a participant had trained. The Moderately Related (MR) condition contained six exemplars from a category that was moderately semantically related to the category on which a participant had trained. The Unrelated (UR) condition contained six exemplars from a category that was semantically unrelated to the category on which a participant had trained.

During transfer, the basic procedure was the same as described for the training phase. Again, 20% of all trials were negative. There were 48 positive trials per condition.

Experiment 1 - Results and Discussion

Mean frame times and accuracies are plotted against training session in Figure 3. Improvement in frame time followed a normal

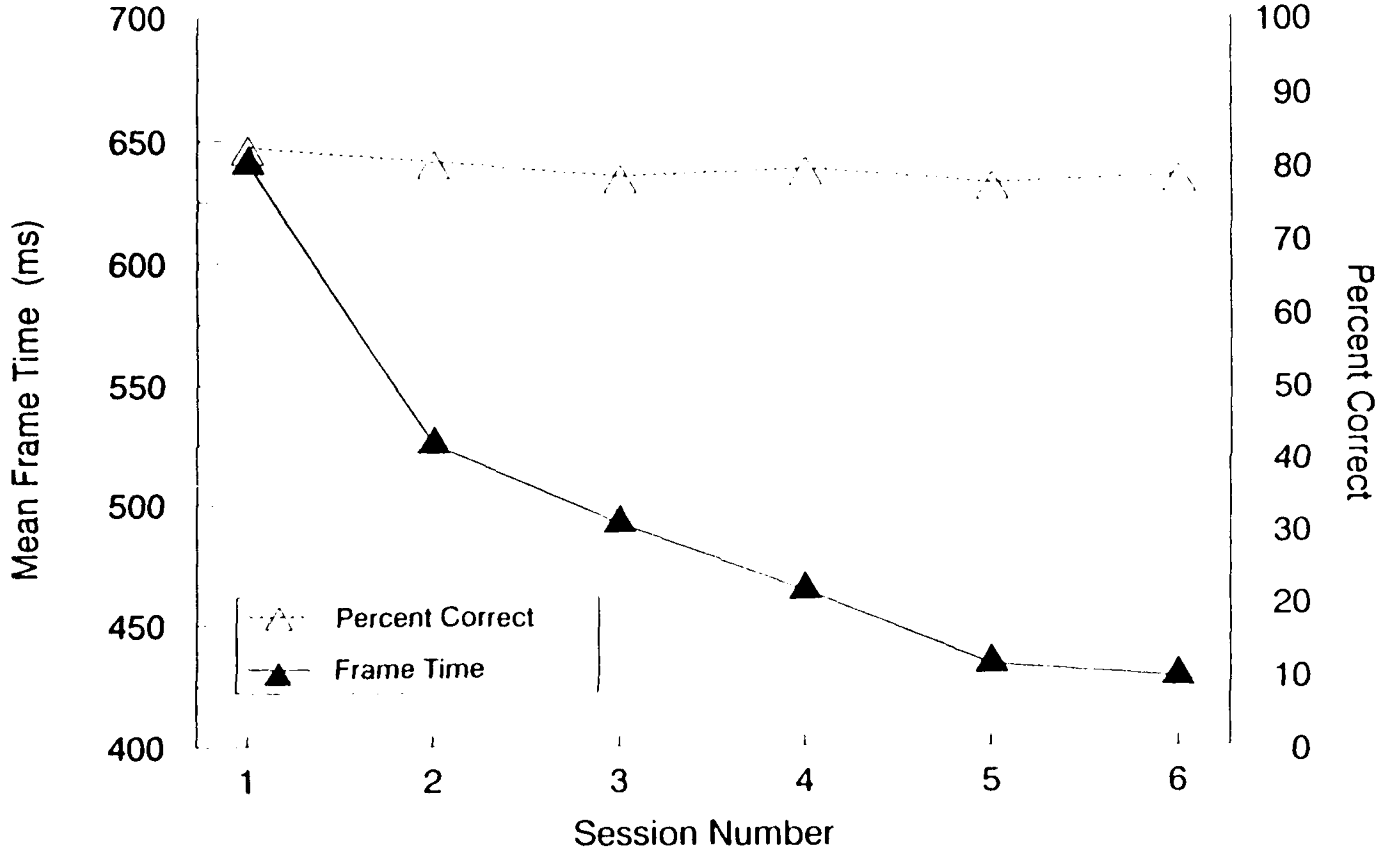


Figure 3. Mean Frame Times and Accuracies are Aggregated Across Participants and Plotted by Session Number. Each point represents 420 trials per participant. Frame time is represented by solid lines and symbols; accuracy, by dashed lines and open symbols.

power function. Accuracy remained relatively stable, ranging from 77% to 80% after the first session.

Accuracy data from the transfer session are plotted by search condition in Figure 4. No clear pattern emerged. A one-way, within-subjects analysis of variance yielded a significant effect of search condition, $F(4,20) = 3.90$, $MS_e = 0.09$, $p < 0.02$. A Newman-Keuls test revealed that none of the untrained conditions were significantly different and the Trained/Trained (T/T) condition was superior only to the Highly Related (HR) and Unrelated (UR) conditions.

Percentage of transfer was measured (cf. Murdock, 1957) by subtracting the control condition (i.e., number of correct positive trials in the UR condition) from the experimental condition (i.e., number of correct positive trials in either the T/T, T/U, HR or MR condition) and dividing the result by the sum of the experimental and control conditions. The result was then multiplied by 100. Although this formula yields smaller percentages than other transfer equations, it is independent of raw score units and yields values at which positive and negative transfer are symmetrical and upper and lower limits are identical. The results are shown in Table 7.

In this experiment, we gave participants a moderate amount of training (2,016 positive trials) on a task driven by visual search. We were interested in whether training on such a task would transfer to new elements which were semantically related to the original task elements. Also, we were interested in whether there would be an ordering of performance based upon the semantic

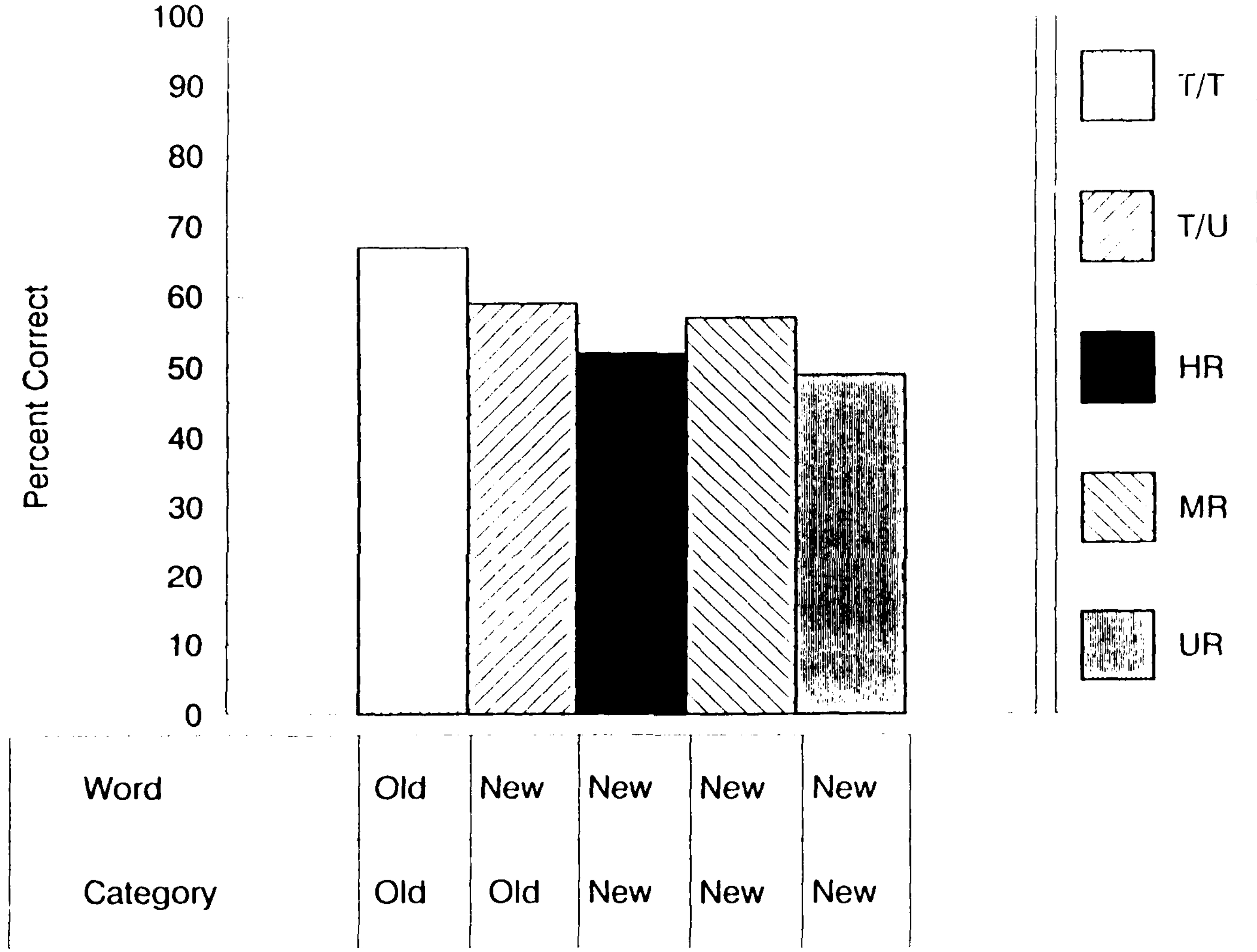


Figure 4. Mean Accuracies (Positive Trials Only) are Aggregated Across Participants and Plotted by Transfer Condition. Each bar represents 48 trials per participant.

Table 7. Percentage of Transfer, Experiment 1

Condition	Transfer Score
Trained/Trained	15%
Trained/Untrained	9%
Highly Related	3%
Moderately Related	8%

relationships between these elements. Although statistically nonsignificant, a trend emerged which indicated that semantic relatedness was influencing performance. First, performance in the three related conditions (T/U, HR and MR) was superior to that for the UR condition. And second, ordering of performance (with the exception of the MR condition) followed a pattern which suggested that semantic relatedness was influencing performance.

Experiment 2 - Overview

In Experiment 2, we tested another group of participants in order to determine whether the minimal amounts of transfer and the nonsignificant differences in performance across conditions found in Experiment 1 were robust phenomena. It was possible that we simply did not provide enough training. Alternatively, it is possible that tasks which are driven by visual search are not conducive to transfer of training, particularly when there is a significant semantic component associated with the task. To this end, we provided participants with an extensive amount of training and "tuned" our experimental paradigm to provide a more rigorous evaluation of these issues.

Experiment 2 - Method

Subjects. Ten right-handed volunteers (5 males, 5 females) were recruited from introductory psychology classes at the Georgia Institute of Technology. Participants were tested for visual acuity of at least 20/30 (uncorrected or corrected) and near vision of at least 20/40. Participants received a combination of research credits and money.

Equipment. All equipment was the same as that described in Experiment 1.

Stimuli. The criteria used to select the categories and exemplars were the same as those described in Experiment 1. The actual stimuli are presented in Appendix C.

Procedure. We used the same "adaptive" multiple-frame procedure described in Experiment 1, with a few modifications. All participants began the experiment at the same "speed," with frame time equal to 850 ms. If a participant's accuracy on any block was equal to or better than 87% correct (26 correct out of a total of 30 trials), frame time on the next block was decreased by 25 ms. If accuracy fell below 77% (23/30), frame time on the next block was increased by 25 ms; otherwise, frame time remained the same. Results from Experiment 1 indicated this allowed accuracy to stabilize around 80% correct. Frame times for an individual's transfer sessions were derived using his or her mean frame time for the final two training sessions. Thus, frame time was held constant during transfer, and accuracy was the dependent measure.

In this experiment, we added the presentation of a Likert-type scale (referred to hereafter as the "certainty scale") to determine a participant's certainty as to whether or not a target was present (see Appendix D for a description of the certainty scale).[2] This scale was presented immediately following the 'T', 'M', 'B' or 'N' response. The participant responded by pressing one of the numeric keys located at the top of the keyboard. The screen then cleared and feedback for that trial was presented (as described in Experiment 1). In addition to the feedback that was presented at the end of each block in Experiment 1, participants received cumulative graphical information about performance on all blocks completed in the current session. Participants could view this feedback as long as they wished. The space bar was pressed when participants were ready to initiate the next block of trials.

Design. All manipulations were within-subject. Data from the following independent variables were collected: (a) position of the target (top, middle, bottom or no target present), (b) frame number of the target exemplar (two through seven), (c) type of trial (positive or negative), (d) number of negative trials in any block (five, six or seven), (e) target category (i.e., memory set), and (f) target exemplar.

The primary dependent variable during training was display set time (speed) and during transfer, correct-incorrect response (accuracy). However, data from the following dependent variables were also collected: (a) certainty scale response (1-5); (b)

---

[2] Use of this scale was prompted by conversations with participants in Experiment 1. As frame times reached the point where participants approached the limits of their perceptual abilities, participants reported difficulty in localizing the target. That is, they claimed that they could see the target (and feedback supported this claim) but that they were uncertain as to the vertical location of the target. The certainty scale provided a technique for examining this phenomenon based on the theory of signal detection. In addition, it provided another metric for examining transfer.

certainty scale response latency; (c) premature responses (on positive trials, responses made before the target exemplar was actually displayed, or on negative trials, before all eight frames were displayed); (d) time spent studying memory set; (e) response latency (on positive trials, the clock started when the target exemplar was displayed and stopped when the response key was pressed; on negative trials the clock started as soon as the last frame was displayed and stopped when the response key was pressed); and (f) type of error (false alarm, miss or error of vertical position).

As in Experiment 1, this study was divided into two phases: training and transfer. Training consisted of one orientation session and 14 training sessions. During the orientation session, we obtained demographic and health information, tested visual acuity and instructed participants on how to perform the task. In addition, participants ran through an abbreviated session -- seven blocks of trials with 30 trials per block, for a total of 210 trials. The 14 training sessions consisted of 14 blocks of trials with 30 trials per block, for a total of 5,880 trials. An average of 20% of all trials were negative (target absent). In any block, five, six or seven negative trials could be presented. The exact number for any particular block was permuted, with the restriction that the mean number of negative trials per block was six.

There were two transfer sessions consisting of 11 blocks each. Five conditions were manipulated across blocks:

1. Trained/Trained (TT)- the same category and exemplars on which an individual had previously trained (three blocks).
2. Trained/Untrained (TU)- six new exemplars from the same category on which a participant had previously trained (two blocks).
3. Highly-Related (HR)- six exemplars from a category which was highly semantically related (see Collen et al., 1975) to the category on which a participant trained (two blocks).
4. Moderately Related (MR)- six exemplars from a category moderately semantically related to the category on which a participant trained (two blocks).
5. Unrelated (UR)- six exemplars from a category unrelated to <u>any</u> other category used in either training or transfer (two blocks).

The five conditions were manipulated between blocks of trials and order of presentation was counterbalanced across participants. Each transfer session for all participants began with one TT block as a "priming" situation (see Appendix E for an example of training and transfer conditions for a typical participant). There were always six negative trials per block.

<u>Experiment 2 - Results and Discussion</u>

<u>Training</u>. Mean frame times and accuracies for each training session are aggregated across participants and presented in Figure 5. By Session 3, accuracy stabilized at about 80%, a level that was expected given the adaptive

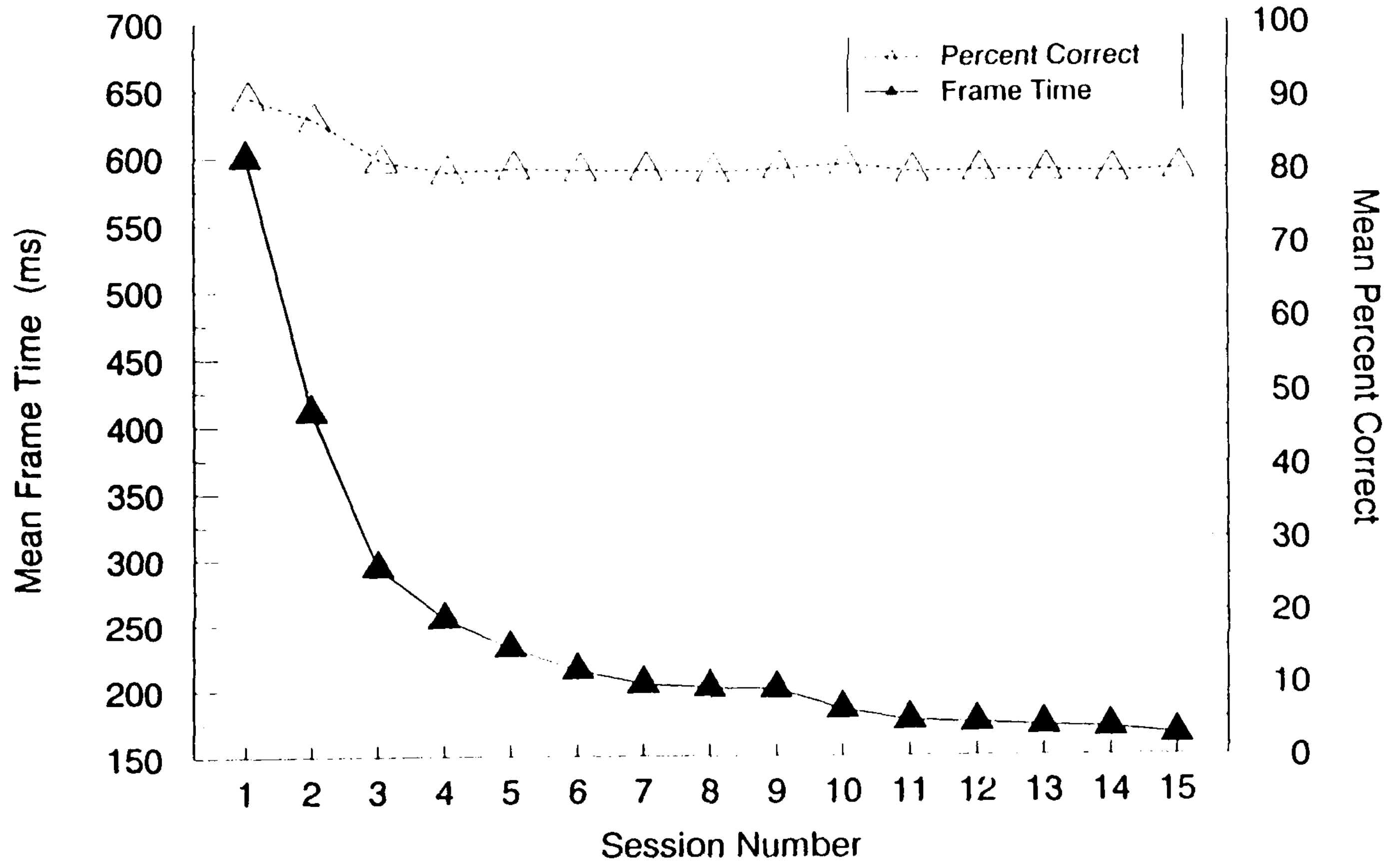


Figure 5. Mean Frame Times and Accuracies are Aggregated Across Participants and Plotted by Session Number. Each point represents 420 trials per participant, except Session 1 (the orientation session) which represents 210 trials. Frame time is represented by solid lines and symbols; accuracy, by dashed lines and open symbols.

training. Frame times decreased according to a normal power function. Following Session 10, improvement was quite modest; participants had become highly skilled at performing the task.

Transfer. The data of main interest pertained to the effects of transfer. Table 8 presents the percentage of transfer for each condition from Sessions 16 and 17. Interestingly, in terms of percentage of transfer, there was no great difference between these findings and those for Experiment 1 of the present investigation. However, the perceptual performance demands in this experiment were considerably more difficult than in Experiment 1 (mean frame time at transfer in Experiment 2 was 160 ms as opposed to 240 ms in Experiment 1). Further examination of Table 8 reveals a distinct pattern consistent with the view that transfer across conditions was influenced by the degree of semantic relatedness to elements of the task on which participants were trained.

Accuracy during transfer was a function of semantic interrelatedness between categories. This is illustrated in Figure 6, which reveals two important findings. First, performance on the T/T condition was superior and approached the 80% level exhibited during training. Second, accuracy on all transfer conditions increased in direct relation to the degree of semantic relatedness to the previously trained category.

Accuracy data from the transfer sessions (Sessions 16 and 17) were aggregated and analyzed with a one-way, within-subjects analysis of variance. There was a significant effect

Table 8. Percentage of Transfer, Experiment 2

Condition	Transfer Score: Session 16	Transfer Score: Session 17
Trained/Trained	20%	16%
Trained/Untrained	10%	8%
Highly Related	7%	5%
Moderately Related	3%	1%

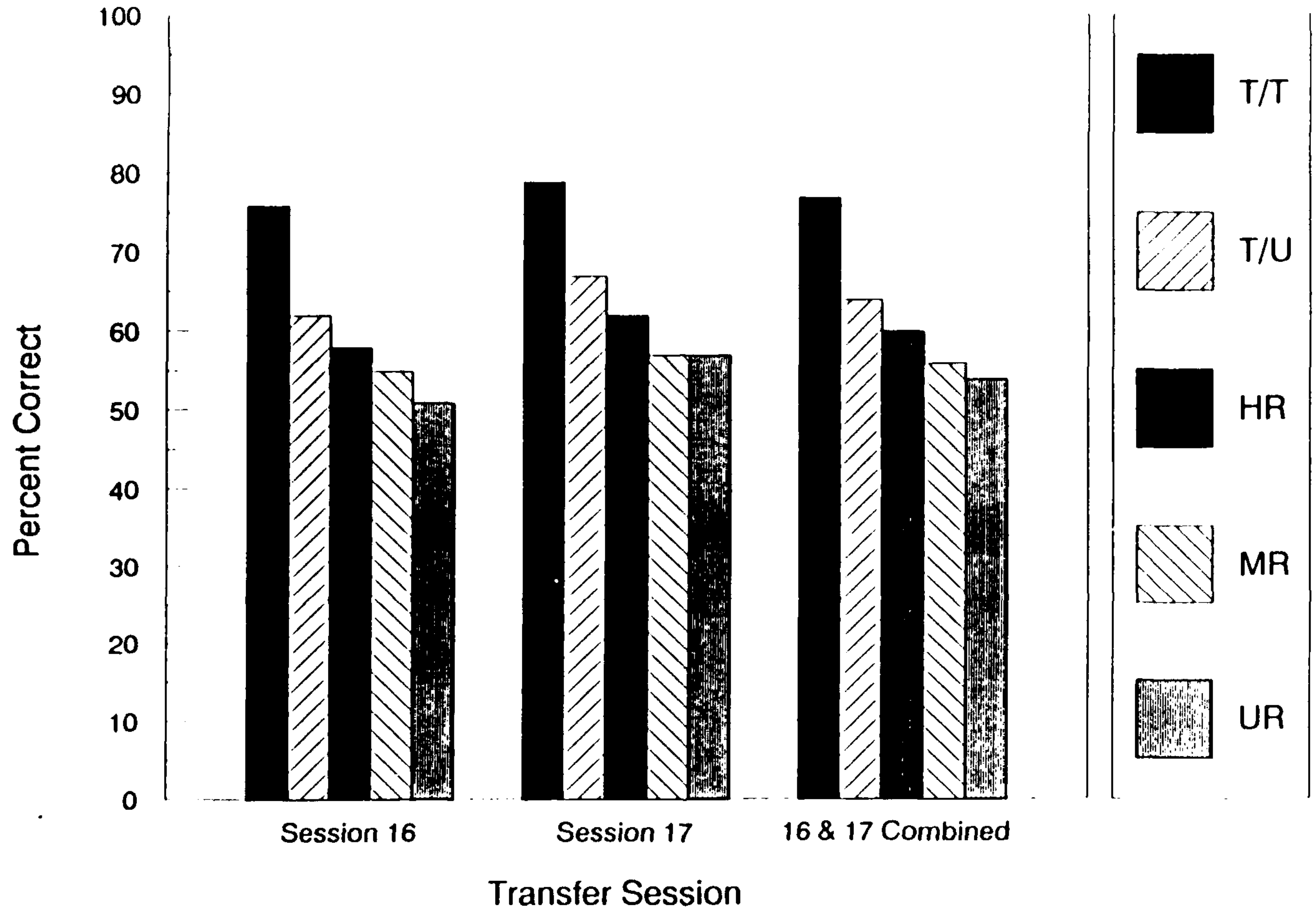


Figure 6. Mean Accuracies (Positive Trials Only) are Aggregated Across Participants and Plotted by Transfer Condition and Session. Each bar represents 48 trials per participant.

of transfer condition, $\underline{F}(4,45) = 18.54$, $\underline{p} < .01$, $MS_e = 0.0186$. A Newman-Keuls test revealed that performance in the T/T condition was superior to that for all other conditions, accuracy was greater for the T/U condition than for both MR and UR, and accuracy in the HR condition was greater than for the UR condition. There was no significant difference between T/U and HR, HR and MR or MR and UR. Thus, for highly related conditions (T/U and HR), we found that performance in the transfer task was superior to that for the control condition (UR).

We found less transfer in visual search than has been reported previously in the memory search literature (e.g., Schneider & Fisk, 1984). However, in our study participants performed at their perceptual processing limits. It is possible that the high degree of overtraining (at extremely brief display durations) induced stimulus feature learning. If so, the amount of transfer may have been attenuated. There is little evidence to suggest that feature learning plays an important role in the transfer of semantically based information. However, if this is the case, and if the brief stimulus displays used in the present study did in fact induce significant feature learning, the amount and pattern of semantic transfer in this study are impressive.

Experiment 3 - Overview

In Experiment 3, we tested a third group of participants in order to obtain another metric of the amount of transfer exhibited in Experiment 2. In this experiment, we tested novice performance under the same conditions in which we tested skilled performers during the transfer phase of Experiment 2.

Experiment 3 - Method

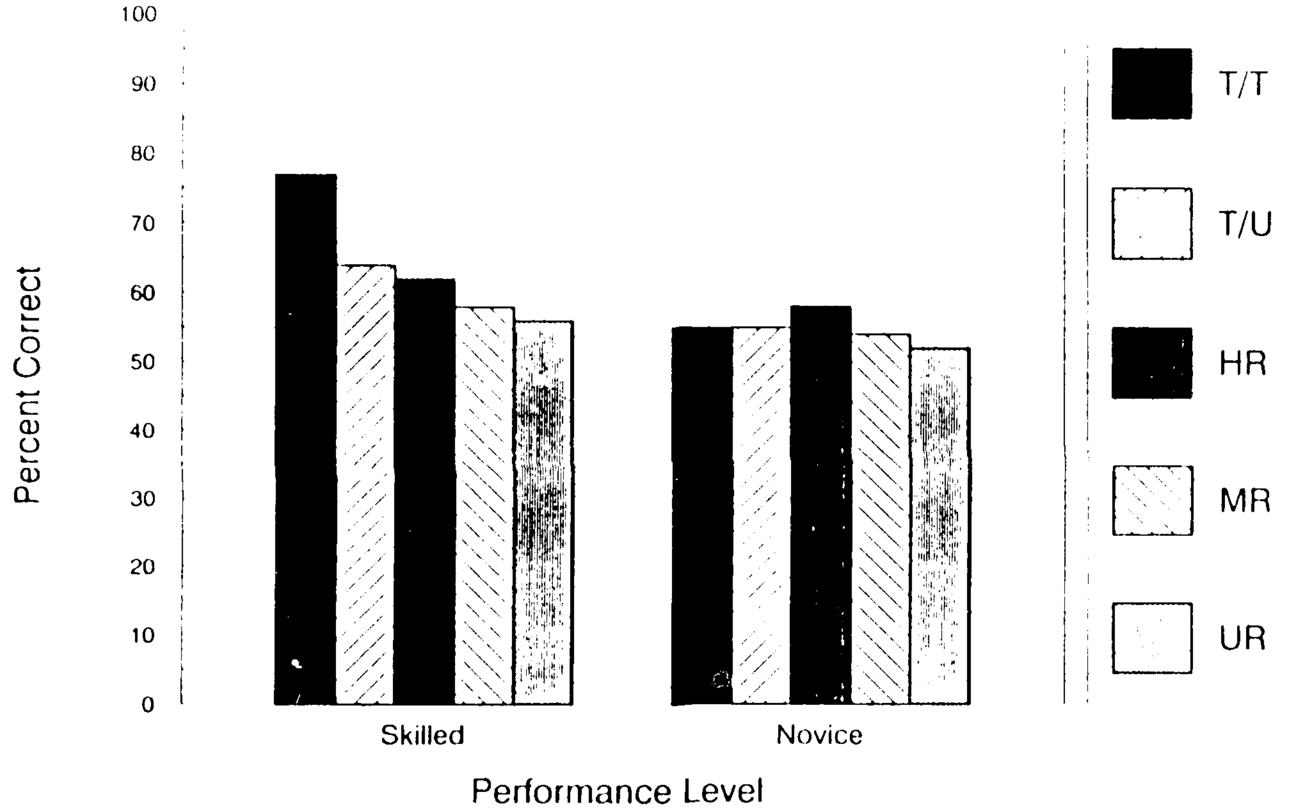
Participants. Ten right-handed volunteers (6 males, 4 females) were recruited from introductory psychology classes at the Georgia Institute of Technology. Participants were tested for visual acuity of at least 20/30 (uncorrected or corrected) and near vision of at least 20/40. Participants received research credit.

Equipment. All equipment was the same as that described in Experiment 1.

Stimuli and Procedure. The stimuli and procedure were the same as those in the transfer phase of Experiment 2, except for the frame time and number of sessions. In this experiment, all participants performed with the same frame time. This frame time was based on the mean frame time for all participants in the transfer phase of Experiment 2, 160 ms. There was only one session.

Experiment 3 - Results and Discussion

Mean accuracies (from all trials) from this experiment and the first transfer session from Experiment 2 are plotted by search condition in Figure 7. As can be seen in the figure, no clear pattern emerged from the novice performers' data and there certainly appears to be no advantage due to semantic relatedness. This finding was also supported statistically. The data from this experiment were analyzed in a one-way, within-subjects analysis of variance. This analysis indicated



<u>Figure 7</u>. Accuracy (All Trials) is Plotted by Performance Level and Transfer Condition (Experiment 2). Skilled performers are from Session 16 in Experiment 2 and novice performers are from Experiment 3. Each bar represents 420 trials per particip. nt.

no statistical differences across search conditions, $\underline{F}(4,36) = 0.56$, $MS_e = 0.01311$, $\underline{p} > 0.6$. Oddly, the best performance among the novice conditions was equal to that of the worst performance among the skilled conditions. In general, novice performance was better than anticipated.

Experimental Series 2 - Summary Discussion

In the present series of studies, we investigated transfer of training in semantic category search tasks in which we examined visual, memory and hybrid visual/memory search. As suggested earlier, the highly speeded performance requirements of the task demanded feature learning. We found that training did indeed transfer positively to highly related components. Although amounts of transfer were modest relative to what has been found in tasks driven by memory search (Schneider & Fisk, 1984), the results are still exciting. As described earlier, given the feature-driven components of the task, the positive influence of semantic relatedness is an important finding.

These results indicate the importance of consistent training in the development of high performance skills dominated by processes associated with visual search. Of particular interest is the finding that consistent training of these skills transfers not only within the same class of stimuli but also to other, highly related stimulus classes. Therefore, one critical component of effective training in such tasks may be training based on highly related examples. Further, these data have significant implications for the development and training of many skills in terms of the level of transfer to be expected in tasks requiring a visual search component.

IV. EXPERIMENTAL SERIES 3: TRANSFER OF AUTOMATIC COMPONENT PROCESSES TO COMPATIBLE, INCOMPATIBLE, AND CONFLICT SITUATIONS -- ISSUES FOR RETRAINING

Introduction

This investigation examined the potential negative effects of developing automatically processed task components when the role of those components changes across tasks. This research allowed a systematic examination of the transfer of automatic processes to situations in which automatic task components were used either in the same way, in an opposite manner, or in conflict with other automatic components. Assessment of performance in the transfer and reversal conditions used in the present study allowed the specification of the deleterious effects of training situations which require either the re-use or the inhibition of previously learned automatic skill components.

Much of the research to date on automatic/controlled processing has focused on the benefits accrued from automatic processing of task components, such as enhancement in the speed and precision of performance as well as a reduction in the amount of resources a given task requires. However, students of skill must be concerned with the potential deleterious effects of incompatible automatic components on learning new skills.

It is a truism (although often overlooked) that automatic processes can have disastrous effects on performance when they are incompatible with task requirements (Norman, 1981; Reason, 1984). Furthermore, these negative effects can persist for quite some time (Shiffrin & Schneider, 1977). Modifying automatic components of skilled performance may demand substantial time on the part of the trainee (and the trainer). Previous findings suggest that retraining to ameliorate the negative influence of a well-learned automatic behavior may require more effort than the initial training of novices (Shiffrin & Schneider, 1977). This fact is important because, when individuals learn complex skills, it is the exception rather than the rule that training focuses on completely novel skill components. Most frequently, trainees bring a repertoire of previously developed skills with them when learning to use new technology or develop new skills.

Little is known about the deleterious effects of competing or incompatible automatic processes on the development of new skills, although this issue has been investigated to some degree in the area of visual search (Dumais, 1979; Fisk & Rogers, 1988; Rogers, 1989). These researchers have examined transfer of automatic process training based on various "strength" models of visual search. In this conceptualization, target and distractor items used in the experimental tasks are hypothesized to have a distribution of strength. Strength may be conceptualized as an item's ability to attract attention (Dumais, 1979; Shiffrin & Czerwinski, 1988; Shiffrin & Dumais, 1981). With practice, items that are attended to consistently (i.e., consistently mapped, or CM targets) become stronger, and items that are consistently ignored (CM distractors) become relatively weaker. A variably mapped (VM) item is, by definition, inconsistent, in that it may appear as a target on one trial and be attended to, but may appear as a distractor on the next trial and therefore be ignored. Consequently, the overall strength of VM items remains the same because they are increased on some trials but decreased on other trials.

The assumption underlying the above research is that there are two component processes in the visual search task which result in automatism: (a) strengthening of consistent targets, and (b) weakening of consistent distractors. Research has demonstrated that if the role of either targets or distractors is reversed (partial reversals - Dumais, 1979; Rogers, 1989) or the roles of both targets and distractors are reversed simultaneously (full reversal - Rogers, 1989; Shiffrin & Schneider, 1977), there will be a disruption in performance. It is proposed that this disruption occurs because participants have difficulty ignoring information that they have been trained to automatically attend to and vice versa. On the other hand, target or distractor components of tasks may be transferred to new task situations with little or no disruption in performance as long as the stimuli continue to serve the same role (i.e., targets continue to be attended to and distractors continue to be ignored).

The preceding summary of research provides evidence that the transfer of automatic components is dependent upon the degree to which the stimuli to be transferred require a similar response, either overt or covert. From a training perspective, all possible recombinations of task components have not been investigated, and many questions remain. Specifically, what will happen to performance if the role of one task component remains the same but the other is reversed within the same condition (christened a "conflict" condition)? The conflict arises, for example, when two CM targets are combined such that the participant must continue to attend to one set but simultaneously learn to ignore the other (previously attended) set. Another open question is how durable the influence of previous training is as a function of the type of component recombination. Though this question has been addressed for full reversal situations (Shiffrin & Schneider, 1977), to our knowledge partial reversal situations have not been examined over time.

The goal of the present experiment was to investigate a range of possible task component recombinations likely to occur in real-world situations: Target Transfer and Distractor Transfer (which should result in little or no disruption), Target Reversal and Distractor Reversal (which should result in significant disruption), and Target Conflict and Distractor Conflict. The two conflict conditions have not previously been tested in the laboratory. Our hypothesis was that both conflict conditions would result in performance disruption; however, the amount of disruption and how it would relate to the disruption found for the partial reversal conditions remained an open question.

Our second goal was to measure the persistence of any disruption that occurred in the transfer conditions. To that end, we provided participants with four sessions of practice on the new conditions. We predicted that the conditions which resulted in the greatest disruption initially would also yield disruption effects of longer duration.

Method

Subjects. Seven male and five female undergraduates participated in this experiment. Students received course credit for up to 4 hours of participation and were paid $5.00 per session for the remainder of their time. All participants were screened for visual acuity of at least 20/30 (far vision) and 20/40 (near vision).

Apparatus. Words from 14 semantic categories were chosen from the taxonomic category norms of Battig and Montague (1969) as stimuli. Categories were screened for hierarchical interrelationships such that no two categories were related (Collen et al., 1975). Eight words from each category, each four to seven letters in length, were chosen as exemplars. Each participant received a unique assignment of categories to each condition.

All stimuli were presented using microcomputers programmed to present the stimuli, collect the responses, and control the timing of the display presentations. The standard keyboard was altered such that the '7', '4', and '1' numeric keypad keys were exchanged with the 'T', 'M', and 'B' keys, respectively. All participants were tested in the same room at individual, partitioned workstations.

Procedure. An individual trial consisted of the following events. A memory set item (category label) was presented to the participant for a maximum of 20 seconds, or until the participant pressed the space bar to initiate the remainder of the trial. Three vertically aligned plus signs were then presented for 0.5

second in the center of the display to localize the subject's gaze. The plus signs were followed by the display set (three category exemplars presented in a column). Participants were given a maximum of 6 seconds to indicate the location of the target (top, middle, or bottom) by pressing the corresponding 'T', 'M', or 'B' key.[3] Participants were encouraged to respond as quickly and accurately as possible.

Participants received feedback after each trial and after each block. After correct trials, response time (RT) was displayed in hundredths of a second. After incorrect trials, a tone sounded and the correct response was displayed. At the end of each block (42 trials), the subject's mean RT and accuracy for that block were presented. If the participants' mean accuracy for any given block fell below 90%, a warning message was displayed, encouraging them to respond more carefully.

Design. The experiment included a training phase and a transfer phase. During the training phase, participants received practice on four consistently mapped (CM) conditions and one variably mapped (VM) condition: CM1 - A(B); CM2 - C(D); CM3 - E(F); CM4 - G(H); VM - IJKLMN(IJKLMN). (Here, for example, the representation A(B) refers to Target Set A displayed with Distractor Set B.) Participants completed ten 1-hour sessions of training, for a total of 8,400 training trials. Each session consisted of 20 blocks (42 trials), five blocks for each training

[3] There was a target present on every trial. Previous research has shown that having a target present on every trial is as effective as having twice as many trials in which half of the trials are target present trials and half are target absent trials in terms of facilitating automatic process development (see Schneider & Fisk, 1980).

condition. The type of training condition was manipulated between blocks and the order of blocks was balanced within each session. In the CM conditions, the target and distractor categories were from distinct sets whereas in the VM condition, the target and distractor sets on a given trial were chosen from a pool of six sets.

In the transfer phase of the experiment, participants participated in four sessions (14 blocks of 42 trials per block), for a total of 2,352 trials. The seven search conditions were: (a) Target Transfer - A(I), (b) Distractor Transfer - J(B), (c) Target Reversal - K(C), (d) Distractor Reversal - D(L), (e) Target Conflict - E(G), (f) Distractor Conflict - F(H), and (g) New CM - M(N). The presentation order of the transfer conditions was counterbalanced across subjects. Each training condition and its corresponding transfer, reversal, or conflict condition is summarized in Table 9.

For the training phase of the experiment, the within-subject independent variable was the Type of Training (CM vs. VM). For the transfer phase, within-subject variables were search condition (the rearrangement of the task components into Target Transfer, Distractor Transfer, Target Reversal, Distractor Reversal, Target Conflict, Distractor Conflict, and New CM) and the number of sessions. The dependent measures were reaction time (RT) and accuracy.

Results

Training Results. Training resulted in standard CM improvement functions. All CM conditions improved in a standard

Table 9. Training and Transfer Conditions

Training	Transfer
A(B)	---> A(I) - Target Transfer
	---> J(B) - Distractor Transfer
C(D)	---> K(C) - Target Reversal
	---> D(L) - Distractor Reversal
E(F)	---> E(G) - Target Conflict
G(H)	---> F(H) - Distractor Conflict
IJKLMN (IJKLMN)	---> M(N) - New CM

Note: The representation A(B), for example, refers to the category set A as the target set and the category set B as the distractor set.

log-log linear fashion, with RT decreasing from about 750 ms to about 560 ms at the end of training (see Table 10 for end-of-training RT data). These decreases were significant, $\underline{F}(9,90) = 45.08$,[4] with no significant interaction among CM conditions and training sessions ($\underline{F} < 1$). Accuracy remained stable for the CM conditions. At the end of practice, none of the CM conditions differed significantly but all were significantly faster than VM $\underline{F}(4,40) = 28.41$.

Transfer Results. The purpose of the study was to examine the effects of different recombinations of CM stimuli on performance. We turn now to those data. We evaluated performance on the transfer conditions relative to performance at the end of training and also relative to the New CM condition (see Table 10 for the transfer performance data). The data presented in Table 10 reflect average performance in the first and last sessions of transfer. What is apparent is that all conditions were slower at transfer relative to the trained CM performance levels. However, relative to the New CM condition (which did not differ from previous VM levels, $\underline{t}(21) = 1.94$) the Target Transfer $\underline{F}(1,60) = 10.82$ and Distractor Transfer $\underline{F}(1,60) = 5.90$ conditions were significantly faster, showing positive transfer. As predicted, the reversal and the conflict conditions resulted in performance disruption, with RT returning to untrained levels. None of these conditions differed from the New CM condition ($\underline{F} < 1$).

---

[4] All analyses reported were significant at an alpha level of .05 unless otherwise stated.

Table 10. Mean RT (in ms) After Training, and in the First and Last (Fourth) Transfer Sessions

After training	Transfer Session 1	Transfer Session 4	
CM - 562	586	555	- Target Transfer
VM - 691	605	540	- Distractor Transfer
	687	638	- Target Reversal
	686	624	- Distractor Reversal
	677	614	- Target Conflict
	644	589	- Distractor Conflict
	662	614	- New CM

As another measure of disruption, we compared performance of the reversal and the conflict conditions relative to the transfer conditions. Both reversal conditions differed from the transfer conditions, $\underline{F}(1,60) = 30.63$. In addition, the Target Conflict condition differed from the Target Transfer condition, $\underline{F}(1,60) = 15.45$. However, perhaps indicative of slightly less disruption, the Distractor Conflict condition did not significantly differ from the Distractor Transfer condition $\underline{F}(1,60) = 2.74$.

Disruption Over Time. Another important issue was how the effects of transfer, conflict, and reversal would persist over time. We compared performance in each transfer condition after four sessions of retraining (see Table 10) to the pre-transfer CM performance level. These comparisons provided information regarding whether or not the retraining compensated for the disruptive effects of the reversal and conflict conditions. After retraining, the two transfer conditions did not significantly differ from pre-transfer CM performance. The Target Reversal $\underline{t}(21) = -4.22$, Distractor Reversal $\underline{t}(21) = -3.58$, New CM $\underline{t}(21) = -3.94$, and Target Conflict $\underline{t}(21) = -2.75$ conditions were all significantly slower than the pre-transfer CM performance level. Importantly, the Distractor Conflict condition did not differ from pre-transfer CM performance.

An analysis of variance conducted on the mean RTs of each condition across all four retraining sessions revealed no interaction with training ($\underline{F} < 1$). This indicated that with the present amount of retraining, equal improvement functions existed for all of the transfer, conflict, and reversal conditions.

These data suggested that, although all conditions were disrupted at transfer, some conditions were more disrupted than others. Given equal improvement functions, only the two transfer and the Distractor Conflict conditions reached the pre-transfer performance level after four sessions of retraining.

Had more training been given, it appears that all transfer conditions would have reached pre-transfer levels. The potential linear function relating performance improvement to a given disruption condition is exciting, as it suggests the potential ability to predict the amount of training required to compensate for disruption at transfer. However, more retraining would be required to determine if improvement for all conditions actually follows a linear function.

Discussion

The present research was undertaken for the purpose of determining how the integration of automatic task components, in conflict with one another, would affect performance. More specifically, we examined situations that required the inhibition of one automatized task component and the re-use of another. The results indicated that competing automatic components disrupted performance to the same extent as the reversal of components. However, it was found that the type of conflicting components affected the severity of that disruption. The Target Conflict and the partial reversal conditions continued to show worse performance than trained CM performance after four sessions of retraining. Distractor Conflict and the transfer conditions were no different than trained CM after only four sessions while the

New CM control condition continued to yield slower performance than trained CM performance.

These data are important from a theoretical perspective because they are thoroughly consistent with the strength models of visual search proposed by Dumais (1979), Shiffrin and Czerwinski (1988), Shiffrin and Dumais (1981), Schneider (1985), and Schneider and Detweiler (1988). For example, in the Target Conflict condition, for any given trial, participants were presented with three high strength words. According to the model, all three stimuli will draw attention with equal strength. In order to detect the target in the Target Conflict situation, the participant must serially scan (use controlled processing) each word in the display because differential strength cannot be used to distinguish the targets from the distractors.

Early in retraining, in the Distractor Conflict condition, the participant is presented with three words that are of low strength. According to the model, because these low strength words repel attention automatically, the only way to search the display and make a response is through controlled processing. However, also according to this model, Distractor Conflict may have improved faster than Target Conflict because participants had only to change the strength for one stimulus. In Target Conflict, however, the strength of two items must be altered. Schneider (1985b) has suggested that CM targets are strengthened more quickly than CM distractors are weakened; hence, the items serving as CM targets in the Distractor Conflict condition could more quickly gain strength relative to the CM distractors.

The present data are important because they describe performance disruption as a function of the type of skill component recombination and retraining of previously acquired automatic component processes. Further, these data provide an approximation of the extent to which these effects are robust even after retraining. The data are of practical value in estimating performance retraining functions in task domains that have benefitted from the direct application of laboratory search/detection results (Carlson et al., 1989: Myers & Fisk, 1987; and see Eggemeier, Fisk, Robbins, & Lawless, 1988, for possible future applications).

This information may be of use to instructional designers in the attempt to incorporate the automatization of task components into their overall training programs. Much has been written concerning the merits of developing automatic components for complex skill acquisition (Ackerman, 1988; Eggemeier et al., 1988; Fisk & Eggemeier, 1988; Schneider, 1985a). Training guidelines have been developed specifying when and how to train consistent task components in a part-task sense for novel tasks. However, transfer and disruption functions should be taken into account when designing training programs, in order to minimize disruption of actual task performance. When planning part-task training of automatic components, it will be important to identify situations in which the automatic components ultimately could be incongruous with whole task demands or with other related tasks. The present data suggest that when identifying consistent task components for automatization, the possible recombinations of those components must also be considered.

V. EXPERIMENTAL SERIES 4: TOWARD AN UNDERSTANDING OF SKILL DECAY -- RETENTION OF AUTOMATIC COMPONENT PROCESSES

Introduction

Typically, investigations of training or skill acquisition are characterized by at least one of the following: examination of the influence of quantity or quality of training and a focus upon a molar level of analysis (i.e., performance of the entire task). Furthermore, there is one common goal in training: to have the operator perform his or her task at a desired level of proficiency a certain amount of time following training. In many instances, the task is to be performed infrequently after long periods of time following training (e.g., scheduled maintenance or emergencies).

As technology has advanced, the role of the worker has evolved considerably. Fewer demands are placed on strength and motor skills of the worker and more on his or her information processing abilities. Due to the intensive effort expended in training individuals to perform complex tasks, the retention of skill over time is an issue of great importance. The majority of research on skill retention has been conducted by or for the military. Typically, these investigations have examined training on, and retention of, tasks involving the assembly or maintenance of equipment (cf. Hagman & Rose, 1983). Typically, these tasks were not analyzed at the task component level; only overall performance was analyzed following some interval of time after training.

According to Fisk and Lloyd (1988), task components must be examined at a molecular level in order to study skill acquisition in terms of internal, stimulus-to-rule relationships. These relationships are intermediate components of the overall task. Without this molecular analysis, it is difficult, if not impossible, to discern task components that are trained readily from components that are trained less readily. Analysis of a skilled task in terms of its individual components is critical if the objective is to identify components that are both stable and transferrable across different tasks and task domains. On an intuitive level, it is clear that an important characteristic of highly skilled performance is its apparent resistance to decay. What is needed is an <u>empirical</u> assessment of the retention of component skills subsequent to development of skilled performance. The present investigation addressed fundamental aspects of the reliability and stability of automatic processing components of skills.

The contributions of automaticity theory to the areas of training and skill acquisition are well documented (Logan, 1985; Schneider, 1985a). According to automaticity theory, skilled performance is driven by two types of information processing: controlled and automatic processing. Controlled processing is characterized as slow, effortful, serial, limited by processing capacity and under direct control of the operator. In contrast, automatic processing is fast, parallel, not constrained by working memory and not under direct control of the operator. Both processes are driven by the consistency of the mapping of the response to stimuli, classes of stimuli or relationships between stimuli (cf. Fisk & Lloyd, 1988).

Regardless of the mapping, controlled processing dominates the earliest stages of training. If major components of the task are consistently mapped (i.e., responses remain the same across practice), automatic processing will come to dominate performance of the task. However, if major components of the task are variably mapped (VM) or inconsistent (i.e., response requirements change across practice), controlled processing will continue to dominate performance.

It is a ubiquitous finding that skill develops only under CM conditions (e.g., Fisk & Schneider, 1983; Myers & Fisk, 1987; and Schneider & Shiffrin, 1977). However, most complex tasks in an applied setting consist of multiple components, some of which may be trained to automaticity, while others remain dependent on controlled processing. To understand the development and maintenance of skilled performance, it is critical to examine a task in terms of these component processes (Logan, 1985; Schneider & Detweiler, 1988).

One area which researchers have not yet investigated is the long-term retention of automatized components of skill. In this study, we provided high, moderate and low amounts of CM training, along with a low amount of VM training, to our participants. Then we measured performance following training at four different intervals across a time span of 180 days. We predicted that performance in the CM conditions would remain superior over VM at each point of measurement following training. We predicted also that there would be a positive relationship between amount of CM training and level of performance on transfer tasks. This investigation was designed to extend understanding of the mechanisms underlying automatic and controlled processing: specifically, the effect of time since training on these processes. Our purpose was to assess differences in the influence of CM and VM training on the retention of skilled performance and to measure the effect of differential amounts of CM training on skill retention.

Experiment 1 - Method

Participants. Twelve volunteers (6 males, 6 females) completed the experiment. Ten were graduate students in psychology at the Georgia Institute of Technology and two were undergraduates. Participants were tested for corrected or uncorrected far vision of at least 20/30 and near vision of at least 20/40, and were paid for their participation.

Equipment. Epson Equity I+ microcomputers equipped with Epson MBM-2095-E monochrome monitors and Epson multimode graphics adapters were programmed to present the task and collect data. The '7', '4' and '1' keys on the numeric keypad were labeled 'T', 'M' and 'B' respectively, to indicate top, middle and bottom (mapping to target positions on the VDU screen). The task was performed within booths constructed of sound-deadening materials, and pink noise was played at a sound level of approximately 55 dB. In this manner, external sounds were masked.

Stimuli. Fifteen semantically unrelated, taxonomic category labels (Collen et al., 1975) from the Battig and Montague category norms (1969) were used as memory-set items in the training and retention phases of the experiment. Ten exemplars from each of these categories were used for display set items, six during training and all ten during retention. Exemplars were selected according to four criteria: visual distinctiveness, semantic distinctiveness, length (between four and seven letters) and prototypicality (highly associated with their respective categories, according to Battig and Montague). A list of these stimuli is presented in Appendix F.

Procedure. Each trial proceeded as follows. The memory set (one, two or three category labels) was displayed in the left center of the video screen at the beginning of each trial. Participants could study the memory set for up to 20 seconds. To view the display set, participants pressed the space bar. An orientation display consisting of three plus (+) signs was presented for 500 ms in the same location as the display set, to allow the participant to focus his or her gaze. Then the display set, consisting of three words in a column, was presented. The participant's task was to identify the target (i.e., an exemplar from one of the categories in the memory set) and to indicate its location (top, middle or bottom) by pressing the corresponding key (labeled 'T', 'M', or 'B') on the keyboard. Participants were allowed a maximum of 6 seconds to enter their responses.

Participants received the following performance feedback. After each correct trial, the participant's reaction time (RT) was displayed. After each incorrect trial, an error tone sounded and the correct response was displayed. Following each block of trials, the participant was given his or her average RT and percent correct for that block. Participants were instructed to maintain an accuracy rate of 95% or better while responding as quickly as possible. If accuracy fell below 90% for any block, the program instructed him or her to respond more carefully.

Design. The experiment consisted of two phases: training and retention. All manipulations in both training and retention were manipulated within-subject and within-block. In the training phase, there were two factors of interest: search condition and memory-set size. Display set size was constant at three. There were four search conditions: (a) high amount of CM training (CM High, 4320 trials), (b) moderate amount of CM training (CM Moderate, 2160 trials), (c) low amount of CM training (CM Low, 720 trials) and (d) VM training (VM, 720 trials). Memory-set size varied from one to three items. There was a target exemplar present on every trial. All manipulations were within-subject and within-block. There were three target categories associated with each CM condition. Six categories were associated with the VM condition; exemplars from these served as both targets and distractors. The six categories associated with the VM condition also served as distractor categories for CM conditions. Assignment of categories to participants was counterbalanced by a partial Latin square. There were 12 sessions lasting an average of 40 minutes each. There were 20 blocks per session and 33 trials per block.

During the retention phase, two new variables were added: exemplar type (trained versus untrained) and retention interval.

In the untrained exemplar condition, four new exemplars were added to each of the categories on which participants had trained previously. The trained exemplar condition was simply the original categories and exemplars on which participants trained. There were five retention intervals: Performance was measured 1 day following training (the criterion retention date, against which subsequent performance was compared), and at 30, 90 and 180 days following training. During each session, participants received 180 trials (60 per condition).

Prior to each retention session (except Day 1), participants received six blocks of CM practice (48 trials per block). This practice took approximately 15 minutes and was provided to allow participants to orient to the experimental environment and task (e.g., practice which keys to press). Categories and exemplars were semantically unrelated (Collen et al., 1975) to those on which participants trained and to those on which they were tested during retention.

The design was a 4 x 3 x 2 x 3 (retention interval x type of training x exemplar type x memory set size) within-subjects factorial design. The dependent variables were accuracy and response latency on correct trials.

Experiment 1 - Results

Training Results. The first key question regarding the training data is: "Did participants become proficient at performing this task?" Mean reaction times and accuracies are presented in Figure 8. In all CM conditions, reaction time (RT) improvement followed a normal power law function, while

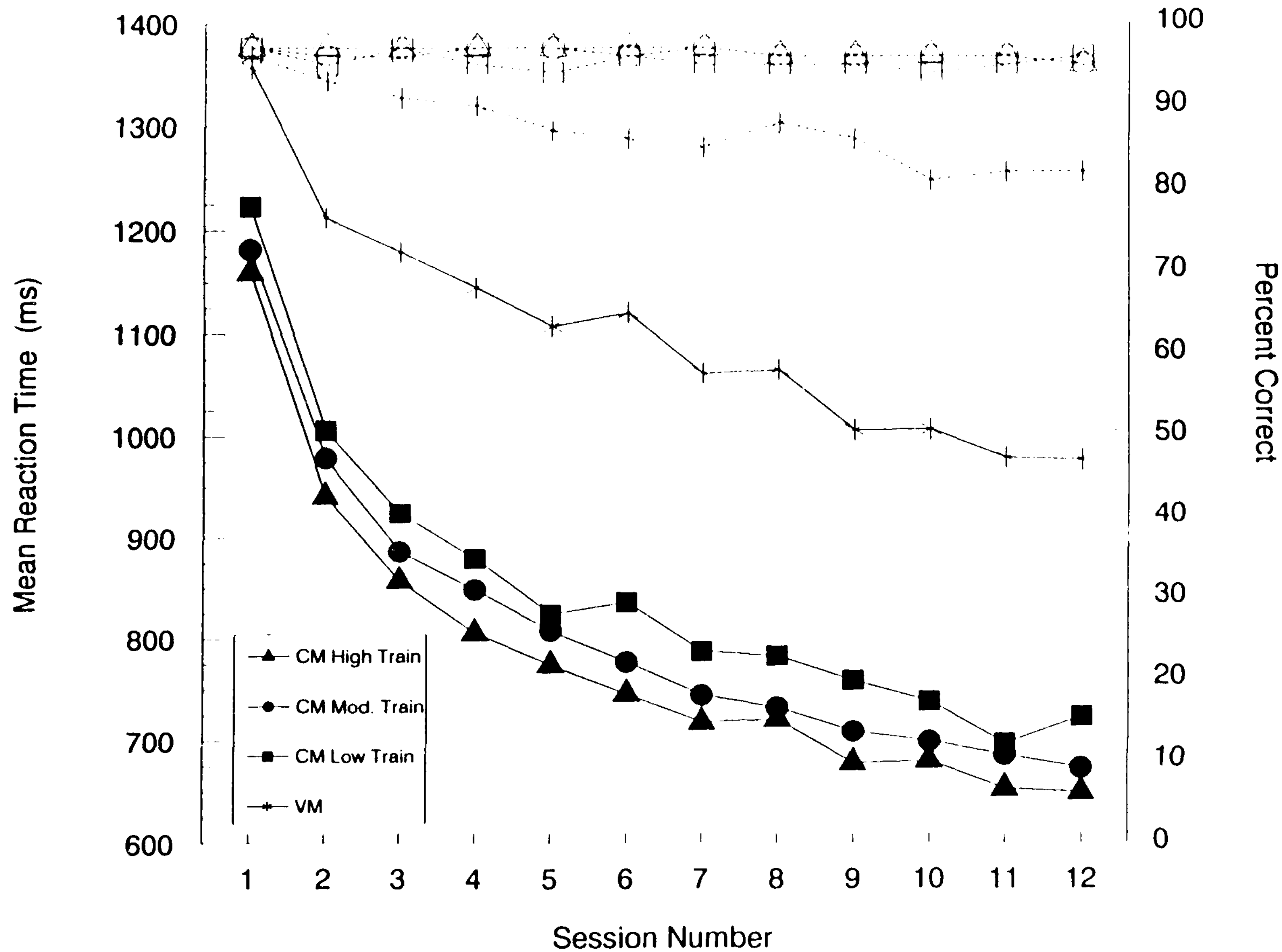


Figure 8. Mean Reaction Times (Correct Trials Only) and Accuracies are Aggregated Across Participants and Plotted by Session Number. Reaction time is represented by solid lines and symbols; accuracy, by dashed lines and open symbols.

accuracies remained stable within the range of 94% to 97% correct. VM RT performance showed modest improvement also. However, VM accuracy declined almost steadily from 96% in Session 1 to 84% by Session 12. This trend demonstrates a classic example of a speed-accuracy trade-off and thus the modest improvement in VM RT performance is not meaningful.

The second key question is: "Are there differences in CM performance due to differential amounts of training?" Comparison of RT means (correct trials only) from Session 12 reveals: The CM High condition was faster (M = 650 ms) than CM Low RT [M= 727 ms, $F(1, 11) = 6.61$, $MSe = 5350$]; CM Moderate was faster (673 ms) than CM Low [$F(1,11) = 5.59$, $MSe = 3041.17$]; and CM Low was faster than VM [M = 1001 ms, $F(1, 11) = 63.09$, $MSe = 6341.35$]. (In all cases the probability level for statistical significance was $p < 0.05$.) RT performance in the CM High condition was slightly faster than CM Moderate but the difference was nonsignificant [$F(1, 11) = 3.10$, $MSe = 1091.61$]. Thus, performance was positively related to amount of CM training.

An examination of comparison slope estimates provides more evidence that increased CM training led to superior performance. These estimates describe the function that relates RT to the number of comparisons (the product of the number of items to be held in memory and the number of items to be searched in the display set) required to make the correct decision. One condition for the attainment of automaticity in visual/memory search tasks is that the slope estimate approach zero (indicating completion of the shift from serial to parallel processing; see

Schneider & Shiffrin, 1977). At Session 12, the comparison slope estimates for CM High, CM Moderate, CM Low and VM were 6.2 ms, 11 ms, 16.6 ms and 53.9 ms, respectively.

Retention Results: Accuracy Data. A 4 x 3 x 2 x 4 (search condition x memory-set size x exemplar training x retention interval) within-subjects analysis of variance was performed on the accuracy data. The main effects of search condition [$F(3,33) = 18.99$, MSe = 0.01[illegible]], memory-set size [$F(2,22) = 49.54$, MSe = 0.123] and exemplar training [$F(1, 11) = 50.62$, MSe = 0.004] were significant. A Newman-Keuls test revealed no differences between CM conditions, but the VM condition was less accurate than any CM condition. There was no effect of retention interval [$F(3,44) = 1.92$, MSe = 0.005], indicating that accuracy across retention intervals was quite stable.

Retention Results: Trained Exemplars. Mean reaction times as a function of retention interval (for all conditions and collapsed across memory set size) are presented in Figure 9. Critical data for this investigation involve the pattern of RT performance decay for trained exemplars across search conditions and retention intervals. A comparison of mean RTs across search conditions reveals that, at Day 1, CM High performance was faster (M = 745.57 ms) than CM Low [M = 813 ms, $F(1,11) = 5.06$, MSe = 5350.13]; CM Moderate was faster (M = 773 ms) than CM Low [$F(1,11) = 5.59$, MSe = 3041.17]; and CM Low was faster than VM [M = 1178 ms, $F(1,11) = 63.09$, MSe = 6341.35]. Following Day 1, there were no statistically significant differences between the

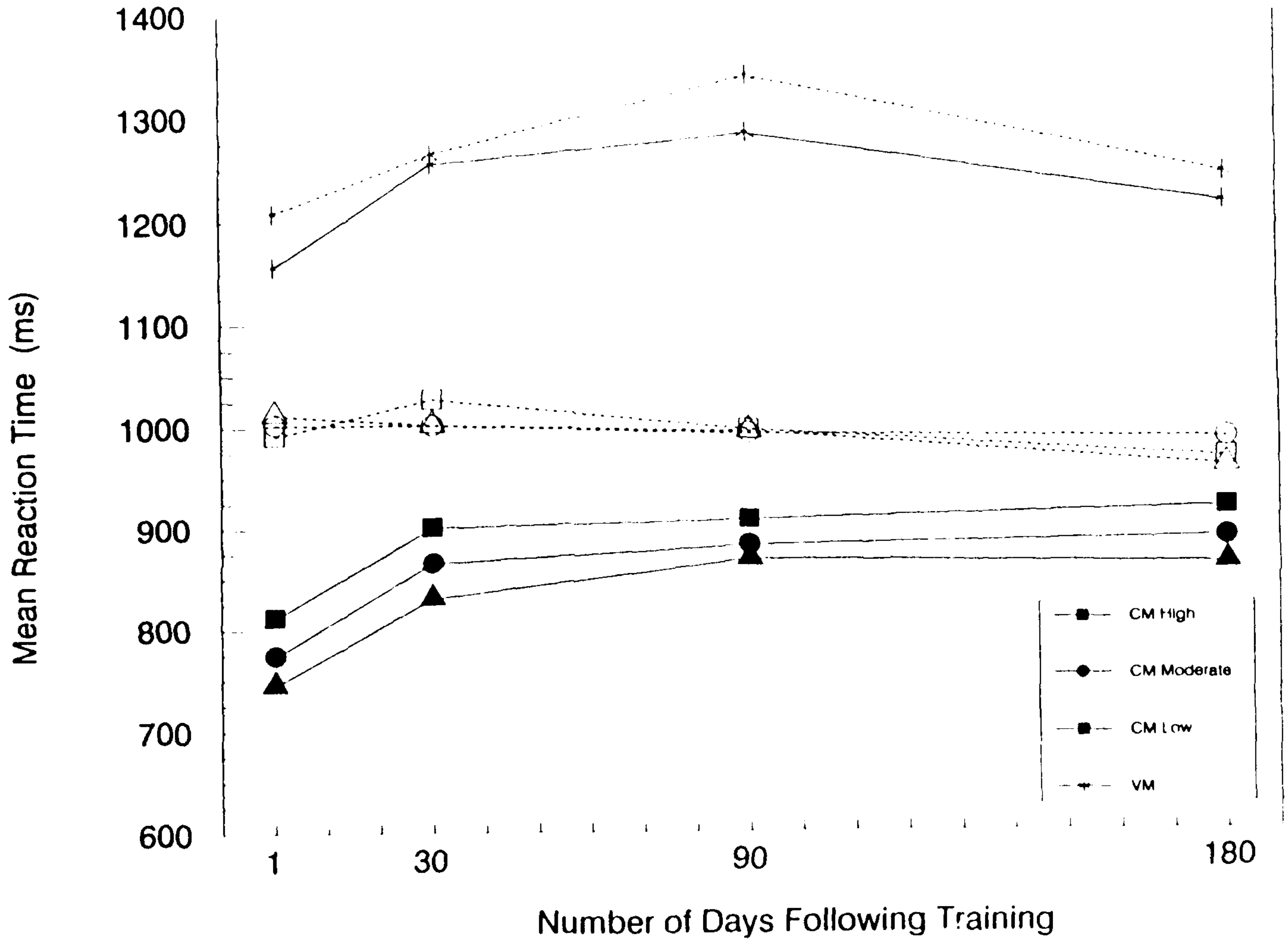


Figure 9. Mean Reaction Times (Correct Trials Only) are Aggregated Across Participants and Plotted by Session Number. Previously trained conditions are represented by solid lines and symbols; untrained conditions, by dashed lines and open symbols.

CM High, Moderate or Low conditions. However, all CM conditions remained superior to VM across all retention sessions.

Comparison of CM RT means across retention intervals reveals that performance in the CM High condition at Day 1 (M = 746 ms) was faster than CM High at Day 30 (M = 830 ms, $\underline{F}(1, 11) = 45.89$, MSe = 927.09]. From Day 30 on, however, performance in the CM conditions did not vary significantly from one retention interval to the next. CM High performance at Day 30 was not significantly different from CM High at Day 90 [M = 869 ms, $\underline{F}(1,11) = 3.33$, MSe = 2718.39]; CM High at Day 90 was not significantly different from CM High at day 180 [M = 865 ms, $\underline{F}(1,11) = 0.09$, MSe = 1898.64]. This pattern also held true for the CM Moderate and CM Low conditions. VM performance was erratic: From Day 1 through Day 90, it decayed, and from Day 90 to Day 180, it improved.

Retention Results: Untrained Exemplars. In the untrained CM search conditions, RT performance was marked by stability. There was even a modest trend of improvement, though not statistically significant. Performance in the untrained VM condition shadowed the trained VM condition (i.e., was not stable). There were no statistically significant differences between the CM High, CM Moderate or CM Low untrained conditions at any retention interval. All untrained CM conditions were superior to both trained and untrained VM conditions across all retention intervals. At Day 1, all trained CM conditions were superior to all untrained CM conditions. With the dramatic decline in performance from Day 1 to Day 30 in the untrained CM conditions, however, the two groups began to converge. Although not statistically significant, continued modest decrement in trained CM conditions combined with modest improvement in untrained CM conditions further reduced differences in performance between the two groups.

Experiment 1 - Discussion

There were four critical results from this experiment: (a) Performance in both CM trained and untrained conditions was superior to all VM conditions at all retention intervals, (b) trained CM conditions exhibited the greatest decrement in performance within 30 days following training; (c) after this initial decline, CM performance remained relatively stable; and (d) the original ordering of performance levels produced by differential amounts of training was maintained at each retention interval.

Interpretation of the pattern of data demonstrated in this study requires consideration of the manner in which automatic processes are developed. Recently, Schneider and Detweiler's (1988) work in formal modelling led them to propose that development of automaticity proceeds in five phases. In order to postulate the development of "pure" automaticity, one needs to examine a range of quantitative and qualitative changes (cf. Shiffrin, 1988). In our investigation, we can point to three pieces of evidence for the development of pure automaticity (certainly in the CM High condition and probably in the CM Moderate condition also): Participants received extensive CM training; improvement in performance followed a power function and had reached asymptote; and final comparison slope estimates approached zero. As noted previously, the CM trained exemplar search conditions exhibited the greatest percentage of decline in RT performance. The largest portion of this decline occurred by Day 30. If we assume that pure automatic processing develops in phases and had developed in these conditions, then the rapid rate at which performance declined from Day 1 to Day 30 suggests that at least some of the underlying mechanisms associated with different phases of automatic processing are somewhat fragile.

After 30 days, RT performance remained quite stable, indicating a limit in the amount of decay of automatic processing. This suggests that the intermediate phases of automaticity are resistant to decay, at least over short intervals of time such as 6 months. This pattern of RT performance -- rapid decline from 1 to 30 days, followed by stable performance from 30 to 180 days after training -- may imply that different processing components are tapped by skilled performance on this task. The task used in this experiment was a hybrid memory/visual search task. Although similar, visual and memory search are dominated by distinct processing mechanisms (Fisk & Rogers, 1989; Flach, 1986). It is possible that retention affects these mechanisms differentially. Experiments 2 and 3 were designed to examine this possibility.

Experiment 2 - Overview

In this experiment, we tested another group of participants in order to examine the effects of retention on a "pure" visual search task. For this task, memory-set size was held constant at one, and display set size was held constant at three.

Experiment 2 - Method

The participants, equipment, stimuli and procedure were the same as those used in the transfer phases of Experiment 2 in Section III of this report (semantic category, visual search task). Participants were tested 30 days following the last transfer session (Session 17).

Experiment 2 - Results and Discussion

In Figure 10, accuracy performance 30 days following the final transfer session (32 days following training) is compared with performance during the final transfer session (Session 17). The condition of interest is the Trained/Trained (T/T) condition. As can be seen, performance in the T/T and Unrelated (UR) conditions exhibited the greatest decline, 7.6% and 10.5%, respectively. A paired comparison of T/T accuracy in Session 17 with T/T accuracy 30 days later (Session 18) reveals that this decline was statistically significant [$\underline{F}(1,40) = 6.01$, MSe = 0.0075]. Performance in the other conditions remained quite stable.

Although the decline in performance was statistically significant, the amount of decline was approximately half that exhibited at Day 30 by the Highly Trained CM condition (Experiment 1 of this series). Certainly, the decline in performance was not sufficient to identify the processing mechanisms associated with visual search as the exclusive locus of the performance decrement exhibited in Experiment 1.

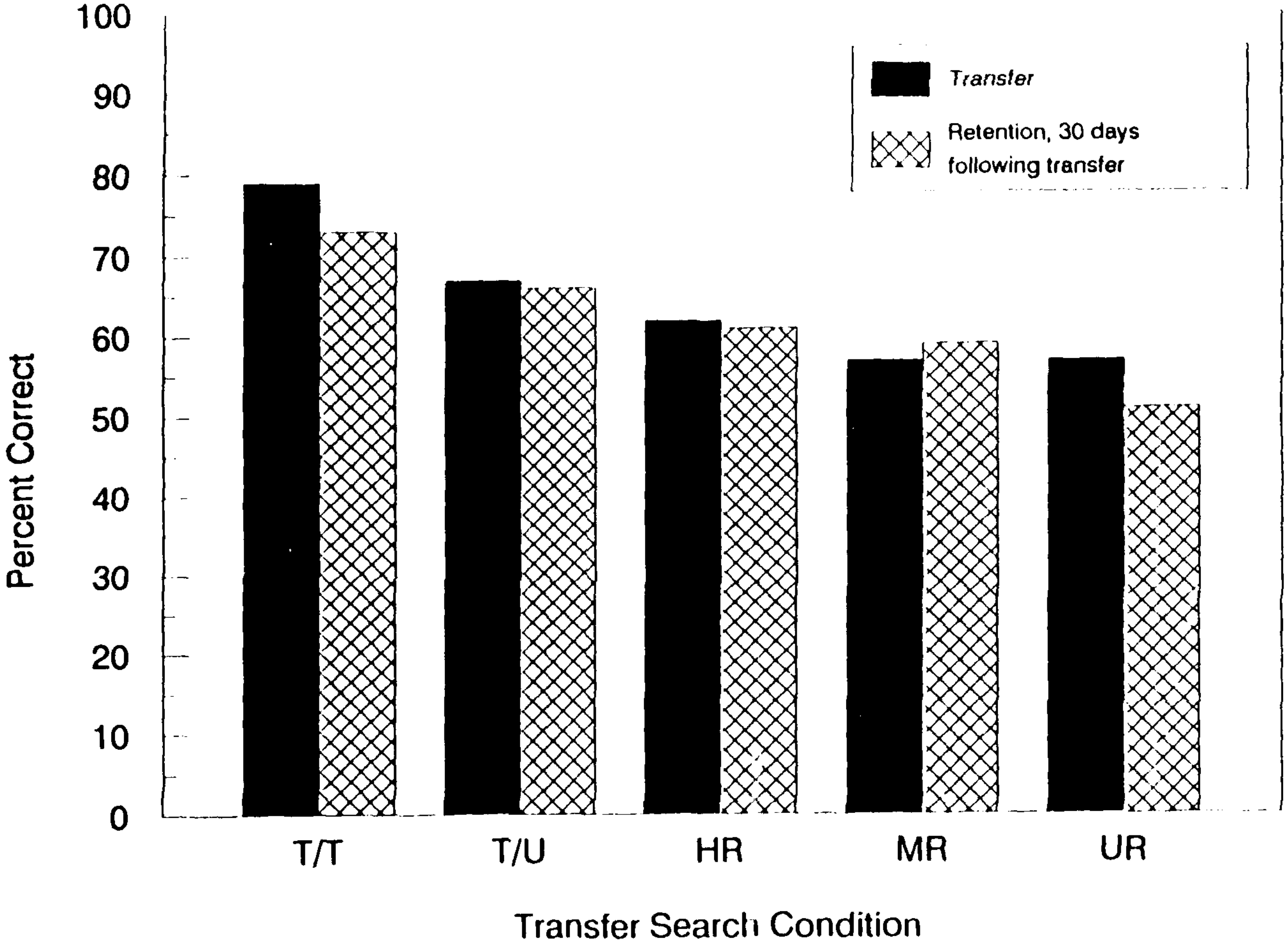


Figure 10. Mean Accuracies (Positive Trials Only) Are Aggregated Across Participants and Plotted by Transfer Search Condition. The graph compares performance at the second day of transfer with performance 30 days later. Each bar represents 48 trials per participant.

Experiment 3 - Overview

In this experiment, we tested another group of participants in order to examine the effects of retention on a "pure" memory search task. For this task, memoiy-set size was varied between one, two, and three items, and display set size was held constant at one item.

Experiment 3 - Method

Participants. Fourteen right-handed volunteers (8 males, 6 females) were recruited from introductory psychology classes at the Georgia Institute of Technology. All participants completed the training phase, but one male and two females failed to return for the retention phase. Participants were tested for visual acuity of at least 20/30 (uncorrected or corrected) and near vision of at least 20/40. Participants received a combination of research credits and money.

Equipment. All equipment was the same as described in Experiment 1, except that the '4' and '5' keys on the numeric keypad were labeled with a 'Y' and an 'N' corresponding to "yes" and "no," respectively.

Stimuli. Fourteen taxonomic category labels (Collen et al., 1975) from the Battig and Montague category norms (1969) were used as memory set items in the training and retention phases of the experiment. Six exemplars from each of these categories were used for display set items during both training and retention. Exemplars were selected according to the criteria described in Experiment 1. These stimuli are presented in Appendix G.

Design. The experiment consisted of two phases: training and retention. All manipulations in both training and retention were manipulated within-subjects and within-blocks. In the training phase, there were two factors of interest: trial type (target present or positive trials versus target absent or negative trials) and memory-set size (1, 2 or 3 category labels). Display set size was constant at one exemplar. Each participant was trained on exemplars from three target categories (six per category) and six distractor categories (six per category). All trials were consistently mapped. Assignment of categories to participants was counterbalanced by a partial Latin square. There were 10 sessions lasting an average of 40 minutes each. There were 19 blocks per session and 42 trials per block, for a total of 7,980 trials of which half were positive trials and half were negative. The retention phase consisted of one session (identical to a training session) conducted 32 days following training.

The design was a 2 x 2 x 3 (phase x trial type x memory set size) within-subjects factorial design. The dependent variables were accuracy and response latency on correct trials.

Procedure. Each trial proceeded as follows. The memory set (one, two or three category labels) was displayed in the left center of the video screen at the beginning of each trial. Participants could study the memory set for up to 20 seconds. To view the display set, participants pressed the space bar. An orientation display consisting of three plus signs was presented for 500 ms in the same location as the display set, to allow the participant to focus his or her gaze. Then the display set, consisting of either one target exemplar or one distractor exemplar, was presented. The participant's task was to decide as quickly as possible whether a target was or was not present and press a key ('Y' for target present and 'N' for target absent) corresponding to his or her decision.

Participants received the following performance feedback. After each correct trial, the participant's reaction time (RT) was displayed. After each incorrect trial, an error tone sounded and the correct response was displayed. Following each block of trials, the participant was given his or her average RT and percent correct for that block. Participants were instructed to maintain an accuracy rate of 95% or better, while responding as quickly as possible. If accuracy fell below 90% for any block, the program instructed him or her to respond more carefully.

Experiment 3 - Results and Discussion

Mean reaction times and accuracies for both target absent and target present trials are plotted against training session in Figure 11. Accuracy data were analyzed with a 2 x 11 x 3 (target presence x session x memory-set size) within-subjects analysis of variance. This analysis revealed no significant differences between the target absent and target present conditions, $F < 1$. Also, accuracies were quite stable; a Newman-Keuls test revealed no statistical differences from Session 4 through Session 11 (the retention session).

During training, reaction times in both conditions followed a normal power function, with RT performance in the target

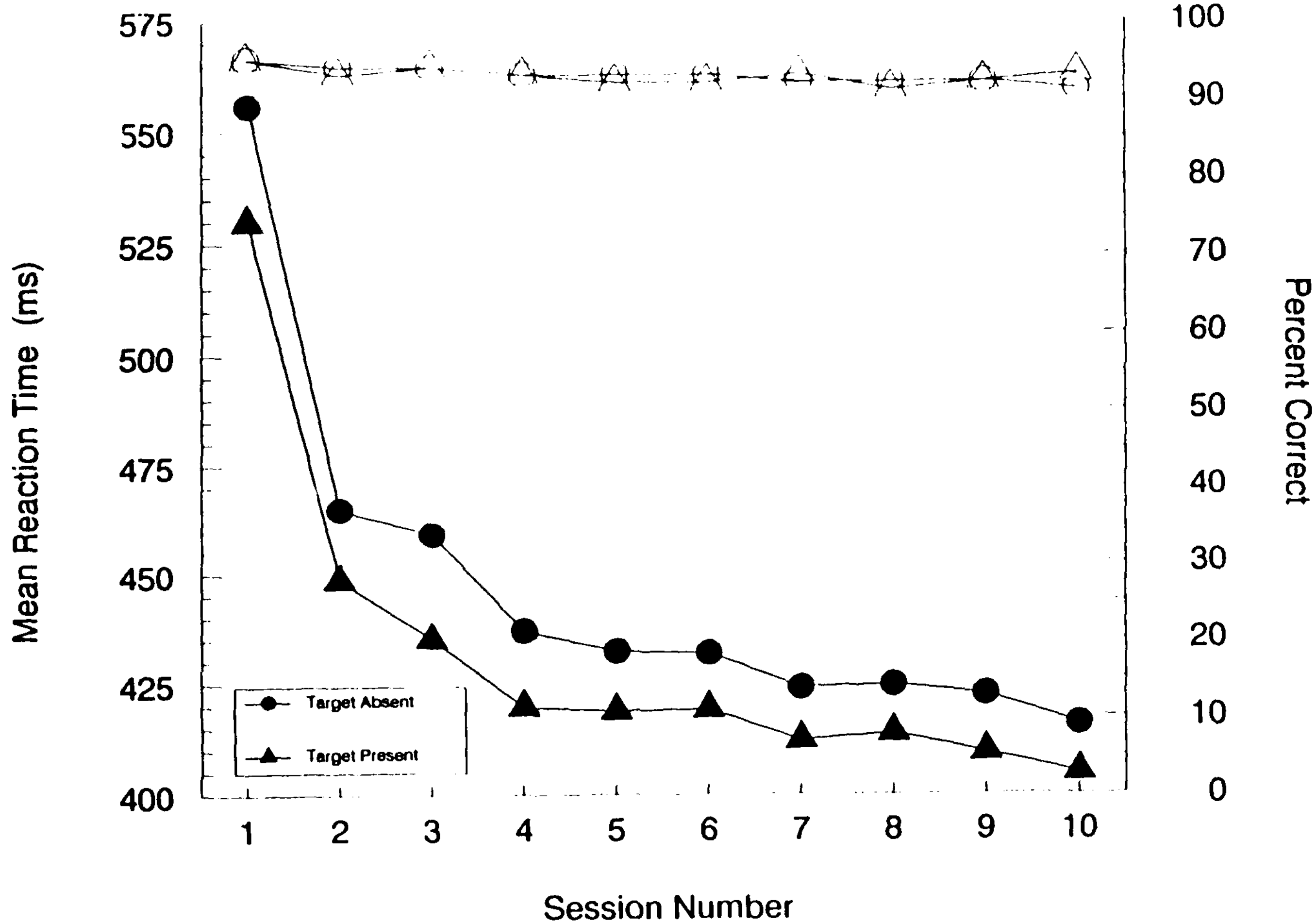


Figure 11. Mean Reaction Times (Correct Trials Only) and Accuracies are Aggregated Across Participants and Plotted by Session Number. Reaction time is represented by solid lines and symbols; accuracy, by dashed lines and open symbols.

present condition maintaining superiority across all sessions. An examination of comparison slope estimates provides more evidence that training led to proficient performance. By Session 8, slope estimates in the target present condition had stabilized at less than 4 ms. After Session 1, slope estimates in the target absent condition never rose above 4 ms.

The central issue, of course, pertains to retention performance: What happened after 32 days without practice? Reaction time performance at the last day of training and 32 days following training (Sessions 10 and 11, respectively) are compared in Figure 12. Clearly, decline in performance was negligible (1.3% in the target absent condition and 1.1% in the target present condition). A paired comparison of RT performance in Sessions 10 and 11 revealed no statistically significant differences [$F(1,10) = 0.77$, $MS_e = 997.11$, $p > 0.4$].

These results indicate that the processing mechanisms associated with memory search are not the locus of performance decrement seen in the first 30 days in Experiment 1. It also appears to be the case that these mechanisms are more resistant to decay than those associated with visual search.

Experimental Series 4 - General Discussion

In the present series of experiments, we examined performance on both automatic and controlled processing component skills at various retention periods from 30 days to 6 months. We explored some of the parameters associated with automatic and controlled processing: in particular, the effect of time since

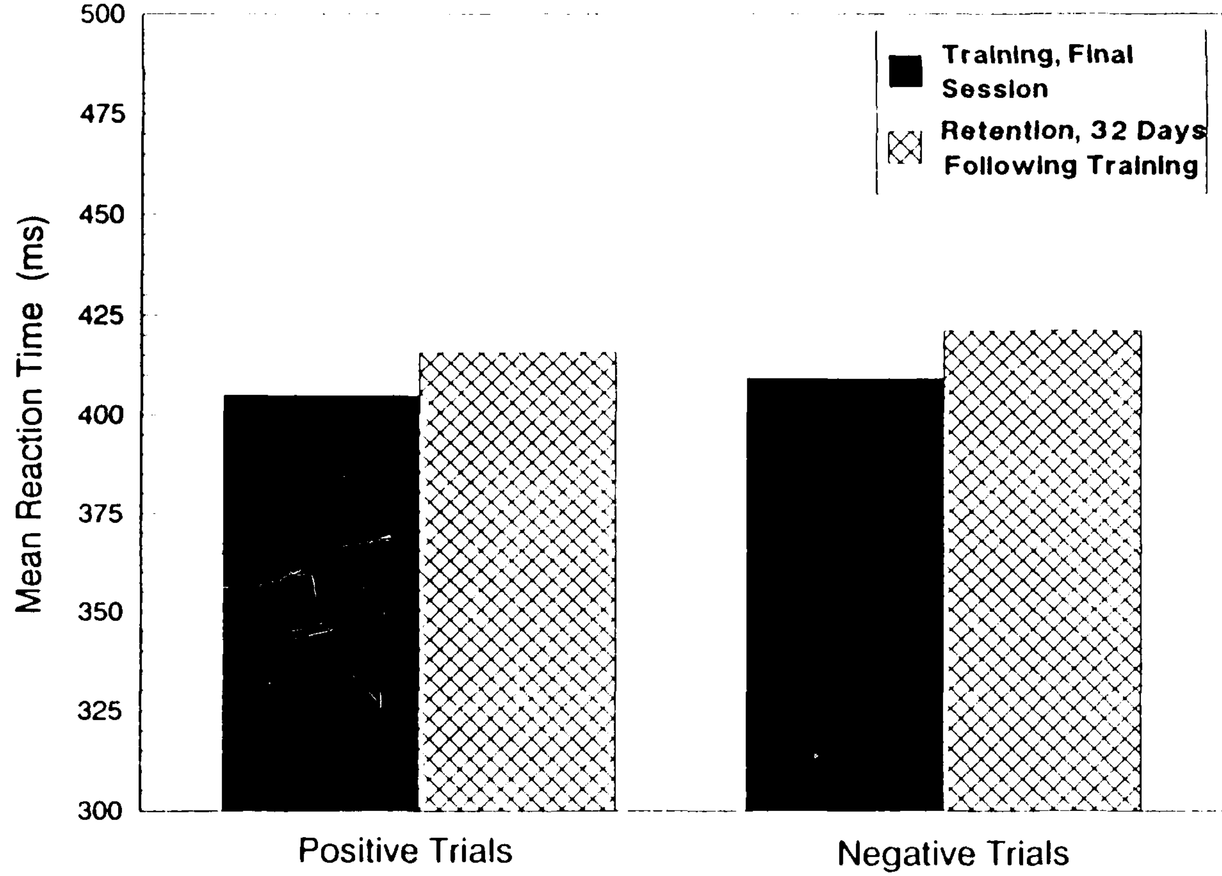


Figure 12. Mean Reaction Times (Correct Trials Only) are Aggregated Across Participants and Plotted by Trial Type. The graph compares performance at the final day of training with performance 32 days later. Each bar represents 399 trials per participant.

training on these processes. Differences in the effect of CM and VM training on the retention of skill were measured, as well as the influence of differential amounts of CM training on skill retention.

The pattern of results demonstrated in these experiments perhaps may be interpreted best within the context of a componential analysis of the processes underlying the complex hybrid memory/visual search task used in Experiment 1. The results of Experiment 3 reveal that access to automatized semantic memory search processes is not disrupted significantly (less than 2%) by an initial retention interval of 32 days. Further, a similar stability of component processes is revealed in Experiment 2, using a visual search paradigm. A performance decrement of less than 8% was demonstrated, a decrement which, although statistically significant, is considerably less than the large diminution in accuracy produced by aggregation of the two task components in the hybrid paradigm of Experiment 1. In the hybrid memory/search task, a decrement of 13% was demonstrated after a 30-day retention interval. The decline in retention performance yielded in the hybrid visual/memory search task cannot be solely attributed to the demonstrated decline in by visual search component nor to that demonstrated by the memory search component. Clearly, there is a qualitative difference between the combined task and the two constituent tasks. What is implicated is an additional degree of complexity present in the hybrid task, a complexity that is absent in either of the individual components.

In the more complex visual/memory search task an increasing level of integration between the mechanisms associated with visual and memory search components may be required. With sufficient CM training, the integration between automatic and controlled processes is facilitated. However, it is possible that periods of inactivity produce an increasing demand upon the integrative mechanism associated with the control structure.

The finding in Experiment 1 that the CM trained exemplar search conditions exhibited the greatest amount of decline in RT performance between Day 1 and Day 30 is noteworthy, as it suggests that it is precisely those skill components that are the most highly overlearned (automatized) that may be most susceptible to decay in the initial period following training. If, in fact, it is the visual search component of the task that declines fastest (in the first 30 days following training), future work may clarify whether or not there is a differential time course for acquisition and decay of this component and for various other components of skilled performance.

The present research has practical and theoretical implications for the study of the effects of time on automatic and controlled processes. The fact that the greatest decay in performance occurred during the first 30 days after training has important implications for designing refresher training strategies to fine-tune automatized components of skilled components.

Given that the decline in performance stabilizes at approximately 30 days following training, it should be possible to predict longer-term performance decrements based upon performance at the 30-day mark. This predictive capability would be invaluable for gauging performance levels across different time spans in a variety of tasks which draw upon both visual and memory search components. The basis for many skilled activities (for example, in a military setting) is to provide training on tasks that remain unused except in emergencies. Identification of the trade-off between amount of training, initial level of performance following training, and level of performance after various periods of delay without practice will allow a more precise assessment of "mission readiness." The present data may also serve to elucidate understanding of the effects of time on skilled performance -- an understanding that is essential to any effort to predict performance after a period of inactivity -- or establish which skill components to emphasize during training.

REFERENCES

Ackerman, P. L. (1988). Determinants of individual differences during skill acquisition: Cognitive abilities and information processing. Journal of Experimental Psychology: General, 117(3), 288-318.

Anderson, J. R. (1982). Acquisition of cognitive skill. Psychological Review, 89, 369-406.

Anderson, J. R. (1983). The architecture of cognition. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1989). Practice, working memory and the ACT* theory of skill acquisition: A comment on Carlson, Sullivan, & Schneider (1989). Journal of Experimental Psychology: Learning, Memory, and Cognition, 15, 527-530.

Battig, W. F., & Montague, W. E. (1969). Category norms for verbal items in 56 categories: A replication and extension of the Connecticut category norms. Journal of Experimental Psychology Monograph, 80(3 Pt. 2), 1-46.

Briggs, G. E. (1969). Transfer of training. In E. A. Bilodeau (Ed.), Principles of skill acquisition. New York: Academic Press.

Bruce, R. W. (1933). Conditions of transfer of training. Journal of Experimental Psychology, 16(3), 343-361.

Carlson, R. A., Sullivan, M. A., & Schneider, W. (1989). Practice and working memory effects in building procedural skill. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15(3), 517-526.

Carlson, R. A., & Yaure, R. G. (1988, November). Random access of component skills in acquisition and problem solving. Paper presented at the annual meeting of the Psychonomic Society, Chicago, IL.

Collen, A., Wickens, D. D., & Daniele, L. (1975). The interrelationship of taxonomic categories. Journal of Experimental Psychology: Human, Learning, and Memory, 1, 629-633.

Cormier, S. M. (1987). The structural processes underlying transfer of training. In S. M. Cormier (Ed.), Transfer of Learning: Contemporary Research and Applications. New York: Academic Press.

Dumais, S. T. (1979). Perceptual learning in automatic detection: Processes and mechanisms. Unpublished doctoral dissertation, Indiana University, Bloomington, IN.

Eggemeier, F. T., Fisk, A. D., Robbins R. J., & Lawless, M. T. (1988). Application of automatic/controlled processing theory to training tactical command and control skills: II. Evaluation of a task analytic methodology. In Proceedings of the Human Factors Society 32nd Annual Meeting, (pp. 1232-1236). Santa Monica, CA: Human Factors Society.

Fisk, A. D., Ackerman, P. L., & Schneider, W. (1987). Automatic and controlled processing theory and its applications to human factors problems. In P. A. Hancock (Ed.), Human factors psychology. New York: North Holland.

Fisk, A. D., & Eboch, M. M. (1987). Applications of automatic/control processing theory to complex tasks: An encouraging look. In Proceedings of the Human Factors Society 31st Annual Meeting, (pp. 674-678). Santa Monica, CA: Human Factors Society.

Fisk, A. D., & Eggemeier, F. T. (1988). Application of automatic/controlled processing theory to training tactical command and control skills: I. Background and task analytic methodology. In Proceedings of the Human Factors Society 32nd Annual Meeting, (pp. 1227-1231). Santa Monica, CA: Human Factors Society.

Fisk, A. D., & Lloyd, S. J. (1988). The role of stimulus rule consistency in learning rapid application of spatial rules. Human Factors, 30, 35-49.

Fisk, A. D., Oransky, N. A., & Skedsvold, P. R. (1988). Examination of the role of "higher-order" consistency in skill development. Human Factors, 30, 567-582.

Fisk, A. D., & Rogers, W. A. (1988). The role of situational context in the development of high performance skills. Human Factors, 30, 703-712.

Fisk, A. D., & Rogers, W. A. (1989). Toward an understanding of age-related memory and visual search effects: Why older adults show an attenuated ability to develop new automatic processes. Unpublished manuscript, Georgia Institute of Technology, Atlanta, GA.

Fisk, A. D., & Schneider, W. (1981). Controlled and automatic processing during tasks requiring sustained attention: A new approach to vigilance. Human Factors, 23, 737-750.

Fisk, A. D., & Schneider, W. (1982). Type of task practice and time-sharing activities predict deficits due to alcohol ingestion. In Proceedings of the Human Factors Society 26th Annual Meeting, (pp. 926-930). Santa Monica, CA: Human Factors Society.

Fisk, A. D., & Schneider, W. (1983). Category and word search: Generalizing search principles to complex processing. Journal of Experimental Psychology: Learning, Memory, and Cognition, 9, 177-195.

Fitts, P., & Posner, M. I. (1967). Human performance. Belmont, CA: Brooks/Cole.

Flach, J. M. (1986). Within-set discriminations in a consistent mapping search task. Perception and Psychophysics, 39, 319-327.

Gick, M. L., & Holyoak, K. J. (1987). The cognitive basis of knowledge transfer. In S. M. Cormier (Ed.), Transfer of learning: Contemporary research and applications. New York: Academic Press.

Gray, W. D., & Orasanu, J. M. (1987). Transfer of cognitive skills. In S. M. Cormier (Ed.), Transfer of learning: Contemporary research and applications. New York: Academic Press.

Hagman, A. M., & Rose, J. M. (1983). Retention of military tasks: A review. Human Factors, 25, 199-213.

Hancock, P. A. (1984). Environmental stressors. In J. S. Warm (Ed.), Sustained attention in human performance. New York: Wiley.

Hancock, P. A., & Pierce, J. O. (1984). Toward an attentional theory of performance under stress: Evidence from studies of vigilance in heat and cold. In A. Mital (Ed.), Trends in ergonomics/human factors I, Amsterdam: North-Holland.

Kramer, A., Schneider, W., Fisk, A. D., & Donchin, E. (1986). The effects of practice and task structure on components of the event related brain potential. Psychophysiology, 23, 33-47.

Kristofferson, M. W. (1977). The effects of practice with one positive set in a memory scanning task can be completely transferred to a different positive set. Memory & Cognition, 5, 177-186.

LaBerge, D., & Samuels, S. J. (1974). Toward a theory of automatic information processing in reading. Cognitive Psychology, 6, 293-323.

Lee, M. D., Rogers, W. A., & Fisk, A. D. (in press). Transfer of automatic component processes to compatible, incompatible, and conflict situations: Issues for retraining. In Proceedings of the Annual Meeting of the Human Factors Society. Santa Monica, CA: Human Factors Society.

Logan, G. D. (1978). Attention in character classification tasks: Evidence for the automaticity of component stages. Journal of Experimental Psychology: General, 107, 32-63.

Logan, G. D. (1979). On the use of a concurrent memory load to measure attention and automaticity. Journal of Experimental Psychology: Human Perception and Performance, 5, 189-207.

Logan, G. D. (1985). Skill and automaticity: Relations, implications and future directions. Canadian Journal of Psychology, 39, 367-386.

Logan, G. D. (1988a). Automaticity, resources, and memory: Theoretical controversies and practical implications. Human Factors, 30, 583-598.

Logan, G. D. (1988b). Toward an instance theory of automatization. Psychological Review, 95, 492-527.

MacKay, D. G. (1982). The problem of flexibility, fluency, and speed-accuracy trade-off in skilled behavior. Psychological Review, 89, 483-506.

Murdock, Jr., B. B. (1957). Transfer designs and formulas. Psychological Bulletin, 54(4), 313-326.

Myers, G. L., & Fisk, A. D. (1987). Training consistent task components: Application of automatic and controlled processing theory to industrial task training. Human Factors, 29, 255-268.

Neches, R., Langley, P., & Klahr, D. (1987). Learning, development, and production systems. In D. Klahr, P. Langley, & R. Neches (Eds.), Production system models of language and development. Cambridge, MA: MIT Press.

Norman, D. A. (1981). Categorization of action slips. Psychological Review, 88, 1-15.

Osgood, C. E. (1949). The similarity paradox in human learning: A resolution. The Psychological Review, 56, 132-143.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), Information processing and cognition. Hillsdale, NJ: Erlbaum Associates.

Prinz, W. (1979). Locus of the effect of specific practice in continuous visual search. Perception and Psychophysics, 25, 137-142.

Rabbitt, P. M. A., Cumming, G., & Vyas, S. M. (1979). An analysis of visual search: Entropy and sequential effects. In S. Dornic (Ed.), Attention and performance VI. Potomac, MD: Erlbaum.

Reason, J. (1984). Lapses of attention in everyday life. In R. Parasuraman & D. R. Davies (Eds.), Varieties of attention (pp. 515-549). Orlando: Academic Press.

Rogers, W. A. (1989). Target and distractor learning in visual search: Age related differences. Unpublished masters thesis, Georgia Institute of Technology, Atlanta, GA.

Rumelhart, D. E., & McClelland, J. L. (1987). Parallel distributed processing: Explorations in the microstructure of cognition (Vol. 1). Cambridge, MA: MIT Press.

Schneider, W. (1985a). Training high performance skills: Fallacies and guidelines. Human Factors, 27, 285-300.

Schneider, W. (1985b). Toward a model of attention and the development of automatic processing. In M. I. Posner & O. S. Martin (Eds.), Attention and performance XI (pp. 475-492). Hillsdale, NJ: Erlbaum.

Schneider, W., & Detweiler, M. (1987). A connectionist/ control architecture for working memory. In G. H. Bower (Ed.), The psychology of learning and motivation, Volume 21. New York: Academic Press, 53-118.

Schneider, W., & Detweiler, M. (1988). The role of practice in dual-task performance: Toward workload modeling in a connectionist/control architecture. Human Factors, 30, 539-566.

Schneider, W., Dumais, S. T., & Shiffrin, R. M. (1984). Automatic and control processing and attention. In R. Parasuraman, & D. R. Davies (Eds.), Varieties of attention (pp. 1-27). Orlando: Academic Press.

Schneider, W., & Fisk, A. D. (1980). Visual search improves with detection searches, declines with nondetection searches. (Report 8004) Champaign, IL: University of Illinois, Human Attention Research Lab.

Schneider, W., & Fisk, A. D. (1984). Automatic category search and its transfer. Journal of Experimental Psychology: Learning, Memory, and Cognition, 10, 1-15.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. Psychological Review, 84, 1-66.

Shiffrin, R. M. (1988). Attention. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), Steven's handbook of experimental psychology (2nd ed., pp 739-811). New York: Wiley.

Shiffrin, R. M., & Czerwinski, M. P. (1988). A model of automatic attention attraction when mapping is partially consistent. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14(3), 562-569.

Shiffrin, R. M., & Dumais, S. T. (1981). The development of automatism. In J. R. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, NJ: Erlbaum.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. Psychological Review, 84, 127-190.

Smith, E. E. (1978). Theories of semantic memory. In W. K. Estes (Ed.), Handbook of learning and cognitive processes, Vol. 5. Potomac, MD: Erlbaum Press.

Sperling, G., Budiansky, J., Spivak, J. G., & Johnson, M. C. (1971). Extremely rapid visual search: The maximum rate of scanning letter for the presence of a numeral. Science, 174, 307-311.

Thorndike, E. L., & Woodworth, R. S. (1901a). The influence of improvement in one mental function upon the efficiency of other functions. (I.) The Psychological Review, 8(3), 247-261.

Thorndike, E. L., & Woodworth, R. S. (1901b). The influence of improvement in one mental function upon the efficiency of other functions. II. The estimation of magnitudes. The Psychological Review, 8(4), 384-395.

Thorndike, E. L., & Woodworth, R. S. (1901c). The influence of improvement in one mental function upon the efficiency of other functions. III. Functions involving attention, observation and discrimination. The Psychological Review, 8(6), 553-564.

APPENDIX A: CATEGORIES AND EXEMPLARS USED IN EXPERIMENT 1 (EXPERIMENTAL SERIES 2)

Training Phase

Target Categories and Exemplars

FRUITS	VEGETABLES
APPLE	CARROT
PRUNES	KALE
ORANGE	BEAN
LIME	RADISH
PEAR	TOMATO
APRICOT	SQUASH
BANANA	CELERY
LEMON	LETTUCE

Distractor Categories and Exemplars

FURNITURE	OCCUPATIONS	WEAPONS
SOFA	NURSE	SWORD
DESK	DOCTOR	PISTOL
TABLE	TEACHER	BOMB
CHAIR	FARMER	CANNON
COUCH	JUDGE	BAYONET
STOOL	CLERK	RIFLE
LAMP	LAWYER	WHIP
DRESSER	DENTIST	KNIFE

BODY PARTS	DWELLINGS	COUNTRIES
BRAIN	HOUSE	CHINA
LIVER	TENT	FRANCE
HEART	CAVE	ENGLAND
LEGS	HOTEL	CANADA
MOUTH	TRAILER	SWEDEN
NOSE	HOME	NORWAY
FOOT	SHACK	JAPAN
HEAD	MANSION	ITALY

Appendix A (Continued)

Transfer Phase

Target Categories and Exemplars

FRUITS (T/T)	VEGETABLES (T/T)	FRUITS (T/U or HR)
APPLE	CARROT	PEACH
PRUNES	KALE	MANGO
ORANGE	BEAN	GRAPE
LIME	RADISH	CHERRY
PEAR	TOMATO	FIGS
APRICOT	SQUASH	PLUM
BANANA	CELERY	
LEMON	LETTUCE	

VEGETABLES (T/U or H/R)	FLOWERS (MR)	CLOTHING (UR)
PEAS	ROSE	SHIRT
ONION	TULIP	PANTS
CABBAGE	DAISY	JACKET
SPINACH	VIOLET	BLOUSE
CORN	ORCHID	DRESS
POTATO	PANSY	SWEATER

NOTE: For example, a participant who trained on vegetables transferred to the following conditions: 1. vegetables (T/T), 2. vegetables (T/U), 3. fruits (HR), 4. flowers and 5. clothing.

Appendix A (Concluded)

Transfer Phase

Distractor Categories and Exemplars

MUSICAL INSTRUMENTS	READING	RELATIVES
TUBA	BOOK	AUNT
CELLO	NOVEL	UNCLE
HARP	PAPER	MOTHER
TRUMPET	JOURNAL	NEPHEW
ORGAN	ARTICLE	SISTER
GUITAR	LETTER	COUSIN
FLUTE	ESSAY	FATHER
FIDDLE	POEM	NIECE

EARTH FORMATIONS	TIME	VEHICLES
RIVER	HOUR	BOAT
ISLAND	WEEK	AUTO
CANYON	MINUTE	SHIP
HILL	YEAR	BICYCLE
CAVE	DECADE	TRAIN
OCEAN	CENTURY	WAGON
VALLEY	SECOND	TAXI
PLATEAU	DAYS	TRUCK

APPENDIX B: PERCENTAGE OF RELATIONSHIP AMONG TARGET CATEGORIES

Categories	Percentage of Relationship
fruits and vegetables	90-99%
fruits and flowers	20-29%
vegetables and flowers	20-29%
fruits and clothing	0%
vegetables and clothing	0%
fruits and musical instruments	0%
vegetables and musical instruments	0%
flowers and clothing	0%
flowers and musical instruments	0%
musical instruments and clothing	0%

APPENDIX C: CATEGORIES AND EXEMPLARS AND USED IN EXPERIMENTS 2 AND 3 (EXPERIMENTAL SERIES 2)

Training Phase

Target Categories and Exemplars

FRUITS	VEGETABLES
RAISIN	PEAS
APPLE	CELERY
PEAR	TOMATO
GRAPE	BEAN
CHERRY	SPINACH
PLUM	SQUASH
PRUNES	ONION
LEMON	RADISH

Distractor Categories and Exemplars

FURNITURE	OCCUPATIONS	TOOLS
SOFA	NURSE	DRILL
DESK	DOCTOR	SANDER
TABLE	TEACHER	RULER
CHAIR	FARMER	WRENCH
COUCH	JUDGE	PLIERS
STOOL	CLERK	LATHE
LAMP	LAWYER	CHISEL
DRESSER	DENTIST	VISE

BODY PARTS	WEAPONS	COUNTRIES
BRAIN	SWORD	CHINA
LIVER	PISTOL	FRANCE
HEART	KNIFE	ENGLAND
LEGS	BOMB	CANADA
MOUTH	CANNON	SWEDEN
NOSE	BAYONET	NORWAY
FOOT	RIFLE	JAPAN
HANDS	WHIP	ITALY

Appendix C (Continued)

Transfer Phase

Target Categories and Exemplars

FRUITS (TT)	VEGETABLES (TT)	FRUITS (TU or HR)
RAISIN	PEAS	FIGS
APPLE	CELERY	ORANGE
PEAR	TOMATO	PEACH
GRAPE	BEAN	BANANA
CHERRY	SPINACH	APRICOT
PLUM	SQUASH	LIME
PRUNES	ONION	
LEMON	RADISH	

VEGETABLES (TU or HR)	FLOWERS (MR)	CLOTHING (UR)
CARROT	ROSE	SOCKS
CABBAGE	LILAC	GLOVES
POTATO	VIOLET	BLOUSE
LETTUCE	DAISY	JACKET
TURNIP	PEONY	DRESS
BEETS	AZALEA	SHORTS

MUSICAL INSTRUMENTS (UR)

TRUMPET
HARP
FLUTE
BANJO
GUITAR
HORN

NOTE: For example, a participant who trains on vegetables will transfer to the following conditions: 1. vegetables (TT) 2. vegetables (TU) 3. fruits (HR) 4. flowers (MR) and 5. either clothing (UR) or musical instruments (UR).

Appendix C (Concluded)

Transfer Phase

Distractor Categories and Exemplars

DWELLINGS	READING	RELATIVES
CASTLE	BOOK	AUNT
TENT	NOVEL	UNCLE
CAVE	PAPER	MOTHER
HOTEL	JOURNAL	NEPHEW
TRAILER	ARTICLE	SISTER
HOME	LETTER	COUSIN
SHACK	ESSAY	FATHER
MANSION	POEM	NIECE

TIME	VEHICLES	EARTH FORMATIONS
HOUR	BOAT	RIVER
WEEK	AUTO	ISLAND
MINUTE	SHIP	CANYON
YEAR	BICYCLE	HILL
DECADE	TRAIN	CAVE
CENTURY	WAGON	OCEAN
SECOND	TAXI	VALLEY
DAYS	TRUCK	PLATEAU

APPENDIX D: RATING SCALE USED IN EXPERIMENTS 2 AND 3 (EXPERIMENTAL SERIES 2)

```
Absolutely Certain                                 Absolutely Certain
NO Target Present          Guess                   A Target Was Present

|-------------|-------------|-------------|-------------|
1             2             3             4             5
```

How certain are you that you that a target was present? (1-5) ==>

APPENDIX E: EXAMPLE OF TRAINING AND TRANSFER CONDITIONS FOR TYPICAL PARTICIPANT (EXPERIMENTAL SERIES 2)

Block Number	Condition	Category	Phase
all blocks	TT	vegetables	training
1	TT (prime)	vegetables	transfer
2 and 7	MR	flowers	transfer
3 and 8	TU	vegetables	transfer
4 and 9	HR	fruits	transfer
5 and 10	UR	clothing	transfer
6 and 11	TT	vegetables	transfer

NOTE: Except for the prime, ordering of presentation of conditions is counterbalanced across participants.

APPENDIX F: CATEGORIES AND EXEMPLARS USED IN EXPERIMENT 1 (EXPERIMENTAL SERIES 4)

Training Phase

Target and Distractor Categories and Exemplars

FURNITURE	ANIMALS	MUSICAL INSTRUMENTS
SOFA	HORSE	TUBA
DESK	TIGER	CELLO
TABLE	BEAR	HARP
CHAIR	RABBIT	PIANO
COUCH	GOAT	FLUTE
STOOL	WOLF	VIOLIN

VEGETABLES	BODY PARTS	WEAPONS
CARROT	BRAIN	SWORD
BEAN	HEART	SPEAR
TOMATO	LEGS	KNIFE
CELERY	HEAD	BOMB
BEETS	NOSE	ARROW
ONIONS	MOUTH	BAYONET

CLOTHING	COUNTRIES	EARTH FORMATIONS
SHIRT	CHINA	RIVER
PANTS	FRANCE	ISLAND
JACKET	ENGLAND	CANYON
GLOVES	CANADA	HILL
SOCKS	SWEDEN	CAVE
DRESS	NORWAY	PLATEAU

Appendix F (Continued)

Training Phase

Target and Distractor Categories and Exemplars

OCCUPATIONS	RELATIVES	BUILDING PARTS
NURSE	AUNT	FLOOR
DOCTOR	UNCLE	CHIMNEY
TEACHER	MOTHER	WINDOW
FARMER	NEPHEW	ROOF
FIREMAN	SISTER	STEPS
JUDGE	COUSIN	CLOSET

VEHICLES	ALCOHOL	UNITS OF TIME
BOAT	BOURBON	HOUR
AUTO	WINE	WEEK
SHIP	VODKA	YEAR
TRAIN	BRANDY	DECADE
TRUCK	SCOTCH	SECOND
BICYCLE	WHISKEY	MINUTE

Appendix F (Concluded)

Retention Phases

Target and Distractor Categories and Exemplars

FURNITURE	ANIMALS	MUSICAL INSTRUMENTS
CABINET	MOUSE	TRUMPET
LAMP	SHEEP	DRUM
DRESSER	LION	OBANJO
BENCH	DONKEY	GUITAR

VEGETABLES	BODY PARTS	WEAPONS
CORN	FOOT	RIFLE
LETTUCE	LIVER	PISTOL
SQUASH	FINGER	TANK
TURNIP	HANDS	CANNON

CLOTHING	COUNTRIES	EARTH FORMATIONS
SHOES	JAPAN	LAKE
SWEATER	ITALY	OCEAN
BLOUSE	SPAIN	VALLEY
SKIRT	RUSSIA	CLIFF

OCCUPATIONS	RELATIVES	BUILDING PARTS
CLERK	FATHER	DOOR
LAWYER	NIECE	WALL
DENTIST	BROTHER	STAIRS
PLUMBER	SONS	CEILING

VEHICLES	ALCOHOL	UNITS OF TIME
TAXI	BEER	CENTURY
JETS	MARTINI	DAYS
JEEP	LIQUEUR	EONS
WAGON	TEQUILA	ERAS

NOTE: The categories and exemplars used in training were also used in retention.

APPENDIX G: CATEGORIES AND EXEMPLARS USED IN EXPERIMENT 3 (EXPERIMENTAL SERIES 4)

Training and Retention Phases

Target and Distractor Categories and Exemplars

FLOWERS	ANIMALS	BODY PARTS
ROSE	HORSE	BRAIN
LILAC	TIGER	HEART
VIOLET	BEAR	LEGS
DAISY	RABBIT	HEAD
PEONY	GOAT	NOSE
AZALEA	WOLF	MOUTH

WEAPONS	EARTH FORMS	MUSICAL INSTRUMENTS
SWORD	RIVER	HARP
SPEAR	ISLAND	FLUTE
KNIFE	CANYON	TRUMPET
BOMB	HILL	BANJO
ARROW	CAVE	GUITAR
BAYONET	PLATEAU	HORN

VEGETABLES	BUILDING PARTS	COUNTRIES
CARROT	FLOOR	CHINA
BEETS	CHIMNEY	FRANCE
CABBAGE	WINDOW	ENGLAND
POTATO	ROOF	CANADA
LETTUCE	STEPS	SWEDEN
TURNIP	CLOSET	NORWAY

Appendix G (Concluded)

Training and Retention Phases

Target and Distractor Categories and Exemplars

CLOTHING	OCCUPATIONS	RELATIVES
SHORTS	NURSE	AUNT
BLOUSE	DOCTOR	UNCLE
JACKET	TEACHER	MOTHER
GLOVES	FARMER	NEPHEW
SOCKS	FIREMAN	SISTER
DRESS	JUDGE	COUSIN

FRUITS	ALCOHOL	UNITS OF TIME
FIGS	HOUR	BOURBON
ORANGE	WEEK	WINE
PEACH	YEAR	VODKA
BANANA	DECADE	BRANDY
APRICOT	SECOND	SCOTCH
LIME	MINUTE	WHISKEY

U. S. GOVERNMENT PRINTING OFFICE: 1990--761-051 20030